Osservazione 7.4.19 *Il sottospazio U di V generato dai seguenti vettori $v, T(v), T^2(v), \dots, T^{m-1}(v)$, è T-invariante.*

Lemma 7.4.20 *Sia U come sopra. Allora esiste un sottospazio complementare di U in V che sia T-invariante, cioè un sottospazio W tale che:*
1) W è T-invariante;
2) $V = U \oplus W$.

Dim. Consideriamo l'insieme $\mathcal{G}$ di tutti i sottospazi G di V che siano T-invarianti e tali che $G \cap U = \{0\}$.

Sicuramente $\{0\}$ verifica queste proprietà, quindi $\mathcal{G}$ non è vuoto. Sia W un sottospazio di V appartenente a $\mathcal{G}$ che sia massimale, cioè tale che non esiste un elemento W di $\mathcal{G}$ che contenga strettamente W. Questo esiste sicuramente, perché altrimenti potrei costruire una catena ascendente di sottospazi infinita, cosa impossibile perché V ha dimensione finita.

Vogliamo dimostrare che $V = U \oplus W$.

Per assurdo, supponiamo che ciò non sia vero. Sia $y \in V$ tale che, però, non sia elemento di $U \oplus W$.

Sia k un intero tale che $T^k(y) \neq 0$, ma $T^{k+1}(y) = 0$. Un tale k esiste sempre perché T è nilpotente. Sia Y il sottospazio di V generato dalla serie: $y, T(y), \dots, T^k(y)$. Chiaramente Y è T-invariante. Quindi anche $W + Y$ è T-invariante.

Se dimostriamo che $U \cap (W + Y) = 0$, abbiamo ottenuto un sottospazio di V, T-invariante e che contiene *strettamente* W (y appartiene a $W + Y$, ma non a W), contro l'ipotesi di massimalità di W.

Dimostriamo allora che $U \cap (W + Y) = 0$.
Per assurdo supponiamo che non lo sia, e quindi esista a U tale che $a \neq 0$ e si scriva come $a = b + g$, con $b \in W$, $g \in Y$ e g$\neq 0$. (Se $g = 0$, $u \in U \cap W$, che è zero per ipotesi). In particolare, $g \in U \oplus W$ e quindi si scrive in modo unico come $g = u + w$. Poiché $g \in Y$, $g = \eta_0 y + \eta_1 T(y) + \dots + \eta_k T^k(y)$.
Sia j il minimo indice per cui $\eta_j \neq 0$. Poiché $T^{k+1}(y) = 0$, mentre $T^k(y) \neq 0$, si ha:

$$T^{k-j}(g) = \eta_j T^k(y) = T^{k-j}(u) + T^{k-j}(w) \,.$$

Quindi $0 \neq T^k(y) \in U \oplus W$. (Ricordiamo che $U \oplus W$ è T-invariante).
Poiché y non appartiene a $U \oplus W$, mentre $T^k(y)$ sì, esiste h intero positivo tale che $x = T^h(y) U \oplus W$, mentre $T(x) = T^k(y) \in U \oplus W$.
Ricordando che U è generato dai vettori $v, T(v), T^2(v), \dots, T^{m-1}(v)$:

$$T(x) = \lambda_0 v + \lambda_1 T(v) + \dots + \lambda_{m-1} T^{m-1}(v) + \tilde{w}, \quad \text{con } w \in W \,.$$

Se $\lambda_0 = 0$, posto

$$\lambda_1 v + \lambda_2 T(v) + \cdots + \lambda_{m-1} T^{m-2}(v) = \tilde{u}\,,$$

si ha $T(x - \tilde{u}) = \tilde{w}$. Quindi $x - \tilde{u} \in W$ (altrimenti $x \in U \oplus W$), ma $T(x-\tilde{u}) = \tilde{w} \in W$. Sia Z il sottospazio di V generato da $x-\tilde{u}$. Il sottospazio $Z + W$ è T-invariante; contiene strettamente W e $(Z + W) \cap U = 0$. (Se fosse $\lambda(x - \tilde{u}) + w = a \in U$, allora $x \in U \oplus W$, cosa non vera). L'esistenza di $Z + W$ contrasta con la massimalità di W.
Se $\lambda_0 \neq 0$, si ha:

$$0 = T^{m-1}(T(x)) = \lambda_0 T^{m-1}(v) + T^{m-1}(w)\,.$$

Siccome $\lambda_0 T^{m-1}(v) \neq 0$, si ha un assurdo perché $\lambda_0 T^{m-1}(v) \in U$ e anche $\lambda_0 T^{m-1}(v) = -T^{m-1}(w) \in W$. Ma per ipotesi $U \cap W = \{0\}$. □

Teorema 7.4.21 *Sia T un endomorfismo nilpotente di uno spazio vettoriale V. Allora $V = U_1 \oplus \cdots \oplus U_h$ dove:*

1) ogni U_i è T-invariante;

2) per ogni U_i esiste un elemento u_i tale che U_i ammette come base una serie del tipo $u_i, T(u_i), \ldots, T^{k_i}(u_i)$;

3) $T^{k_i+1}(u_i) = 0$, cioè $T_{|U_i}$ è nilpotente di ordine $k_i + 1$.

Dim. Basta applicare reiteratamente il Lemma 7.4.20 al sottospazio W, tenendo presente che $T_{|W}$ è nilpotente di ordine $m' \leq m$. □

Osservazione 7.4.22 *Se il sottospazio U è generato dalla serie di vettori $T^m(u), T^{m-1}(u), \ldots, T(u), u$, e la scegliamo, in quest'ordine, come base di U, allora l'endomorfismo nilpotente T sarà espresso in questa base dalla matrice:*

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}. \tag{7.10}$$

Infatti, $T(T^m(u)) = 0$, *mentre, per gli altri elementi della base, poiché* $T(T^i(u)) = T^{i+1}(u)$, *l'immagine è l'elemento immediatamente precedente.*

La matrice A *è una matrice che ha tutti zeri tranne che nei posti* $(i, i+1)$ *immediatamente sopra alla diagonale principale.*

Se T è un endomorfismo nilpotente di V, poiché $V = U_1 \oplus \cdots \oplus U_h$, con U_i come nel Teorema 7.4.17, allora scegliendo come base di V l'unione delle basi degli U_i come nell'Osservazione 7.10, la matrice associata a T è :

$$M = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & A_h \end{pmatrix},$$

dove ciascun A_i è del tipo (7.10) ed è di ordine uguale alla dimensione di U_i, quindi decrescente (per l'osservazione contenuta nella dimostrazione del Teorema 7.4.21).

Osservazione 7.4.23 *La matrice* M' *che si ottiene permutando i blocchi* A_i *di* M, *è simile a* M. *Infatti la disposizione dei blocchi in* M *dipende dall'ordine con cui vengono prese le basi degli* U_i *per andare a formare la base di* V. *Scegliendo un diverso ordine, i blocchi vengono permutati.*

Definizione 7.4.24 *La matrice quadrata di ordine* k

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \lambda & 1 \\ 0 & \cdots & \cdots & \cdots & \lambda \end{pmatrix} = A + \lambda I_k \,, \qquad (7.11)$$

con A *come in (7.10) e* $\lambda \in \mathbb{R}$, *si dice* blocco di Jordan *di ordine* k *relativo a* λ.

Osservazione 7.4.25 *Per l'Osservazione 7.4.17, se* F *è un endomorfismo con un solo autovalore* λ, $T = F - \lambda I$ *è nilpotente.*

Siccome in una certa base la matrice associata a T è una matrice diagonale a blocchi M, con blocchi come in (7.10), nella stessa base la matrice associata a F è $M + \lambda I_k$, ossia del tipo:

$$J(\lambda) = \begin{pmatrix} J_{k_1}(\lambda) & 0 & \cdots & 0 \\ 0 & J_{k_2}(\lambda) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & J_{k_h}(\lambda) \end{pmatrix}, \tag{7.12}$$

cioè una matrice diagonale a blocchi di Jordan di ordine non crescente (cioè $k_i \geq k_i + 1$) relativi all'autovalore λ.

Teorema 7.4.26 (Jordan) *Sia data una matrice A con tutti gli autovalori $\lambda_1, \ldots, \lambda_h$ nel campo degli scalari. Allora A è simile ad una matrice J_A della forma:*

$$J_A = \begin{pmatrix} J(\lambda_1) & 0 & \cdots & 0 \\ 0 & J(\lambda_2) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & J(\lambda_h) \end{pmatrix}, \tag{7.13}$$

dove $J(\lambda_i)$ è una matrice di ordine uguale alla molteplicità algebrica di λ_i, di tipo (7.12), cioè diagonale a blocchi di Jordan relativi allo autovalore λ_i.

La matrice J_A si dice *forma canonica di Jordan* di A.

Enfatizziamo il diverso significato di tutte queste matrici J.

a) $J_k(\lambda)$ è il blocco di Jordan elementare: tutti λ sulla diagonale principale, tutti 1 sulla sopradiagonale e poi tutti 0.

b) $J(\lambda)$ è una matrice diagonale a blocchi che ha vari blocchi di Jordan sulla diagonale ma *tutti relativi allo stesso autovalore* λ.

c) J_A è una matrice diagonale a blocchi che ha sulla diagonale matrici del tipo precedente $J(\lambda)$, relative però a tutti i vari autovalori di A.

La dimostrazione del Teorema 7.4.26 è in sostanza tutta la trattazione qui svolta. Ricapitoliamola.

1) Data una matrice A con tutti gli autovalori nel campo base. Si consideri

l'endomorfismo F che A rappresenta in una certa base. Gli autovalori di F sono gli stessi di A, tutti appartenenti al campo degli scalari e pertanto possiamo decomporre V in somma diretta degli autospazi generalizzati relativi a ciascun autovalore:

$$V = W(\lambda_1) \oplus \cdots \oplus W(\lambda_h)\,.$$

2) Poiché i sottospazi $W(\lambda_i)$ sono F-invarianti, la matrice di F, in una base di V che sia l'unione delle basi dei singoli $W(\lambda_i)$, è una matrice diagonale a blocchi, dove ciascun blocco è la matrice associata a F_i, F ristretto a $W(\lambda_i)$, nella base scelta.
3) F_i ha un solo autovalore, λ_i, su $W(\lambda_i)$, quindi $F - \lambda_i I$ è nilpotente.
4) $W(F_i)$ viene decomposto in una somma diretta di sottospazi U_r, F-invarianti, ciascuno dei quali è generato da una serie:

$$T^m(u_r), T^{m-1}(u_r), \ldots, T(u_r), u_r, \quad \text{dove } T = F - \lambda_i I\,. \tag{7.14}$$

4) Su ciascun U_r, F è espresso da un blocco di Jordan relativo a λ_i di ordine $m = \dim U_r$, se si sceglie come base di U_r la serie (7.14).
5) Su ciascun $W(\lambda_i)$, F è espresso da una matrice del tipo (7.12), $J(\lambda_i)$, che è una matrice di ordine $m_a(\lambda_i)$ diagonale a blocchi, fatta da blocchi di Jordan di ordine vario, ma tutti relativi allo stesso autovalore λ_i, se si sceglie come base su $W(\lambda_i)$ l'unione delle basi scelte per gli U_r.
6) Su tutto V la matrice associata a F sarà una matrice diagonale a blocchi, ciascun blocco essendo la matrice di cui al punto precedente, se su V si sceglie come base l'unione delle basi dei $W(\lambda_i)$. Questa matrice è la forma di Jordan di F e la base di V per cui questo accade è detta base di Jordan.
7) Questa matrice e A sono simili perché rappresentano lo stesso endomorfismo in basi diverse.

Resta da dimostrare la sua *unicità* una volta stabilito l'ordine in cui vengono considerati gli autovalori di F (ad esempio in ordine crescente).

Dal procedimento precedente segue che l'ordine delle varie matrici $J(\lambda_i)$ è fissato da A essendo uguale a $m_a(\lambda_i)$.

Quando si affronta la scomposizione di $W(\lambda_i)$, la dimensione di U_1, che determinava l'ordine del primo blocco di Jordan, è data dall'ordine di nilpotenza di $T = F - \lambda_i I$ ristretto a $W(\lambda_i)$, che determinava la lunghezza della prima serie. La dimensione del blocco di Jordan successivo è dunque dato dall'ordine di nilpotenza di T ristretto a W, un complementare di U_1

T-invariante. Questo W non è univocamente determinato. Potrebbero esistere due diversi sottospazi di $W(\lambda_i)$, W_1 e W_2, tali che siano T-invarianti e $W(\lambda_i) = U_1 \oplus W_1 = U_1 \oplus W_2$. Ma se accade che l'ordine di nilpotenza di T ristretto a W_1 è uguale all'ordine di nilpotenza di T ristretto a W_2 abbiamo finito.

Siano allora h e k gli ordini di nilpotenza di T ristretto rispettivamente a W_1 e W_2 e supponiamo per assurdo che $h < k$. Esiste quindi $w_2 \in W_2$ tale che $T^h(w_2) \neq 0$. Inoltre $T^h(w_2) \in W_2$ perché questo è T-invariante. Poiché $W(\lambda_i) = U_1 \oplus W_1$, allora $w_2 = u + w_1$ (con ovvio significato). Abbiamo $T^h(w_2) = T^h(u) + T^h(w_1)$. Ma per definizione di h, $T^h(w_1) = 0$ e dunque: $0 \neq T^h(w_2) = T^h(u) \in U_1$, mentre per ipotesi $U_1 \cap W_2 = 0$.

Teorema 7.4.27 *Due matrici sono simili se e solo se $J_A = J_B$, cioè se e solo se hanno la stessa forma di Jordan (a meno di permutazione dei blocchi).*

Dim. Siccome A è simile a J_A e B è simile a J_B, se $J_A = J_B$, per la proprietà transitiva della similitudine A è simile a B.

Viceversa, se A e B sono simili, esse rappresentano lo stesso endomorfismo F su V e pertanto, siccome F, in una base di Jordan, è associato alla matrice J, A e B hanno la stessa forma di Jordan J. □

Dalle costruzioni utilizzate per le dimostrazioni si può estrapolare un metodo per ricavare, data una matrice A, la sua forma canonica e una base di Jordan.

Per comodità riportiamo a seguire lo schema di queste costruzioni e sarà utilissimo esercizio ritrovare nella teoria sovraesposta la giustificazione dei passi dello schema stesso.

Ricetta per "Jordanizzare" una matrice. Sia A una matrice quadrata di ordine n.

1) Determinare il polinomio caratteristico $p_A(t)$ di A.

2) Determinare tutte le radici di $p_A(t)$, cioè gli autovalori di A, con la loro molteplicità algebrica. Siano esse $\lambda_1, \dots, \lambda_h$ e $ma(\lambda_i)$ sia la molteplicità algebrica di λ_i.

   Se i λ_i sono tutti reali, bene. Altrimenti si può procedere ugualmente, ma la forma di Jordan di A sarà una matrice con valori complessi λ nei vari blocchi.

   Per la definizione della forma di Jordan reale per matrici con autovalori complessi rimandiamo al paragrafo successivo.

3) Si prende in esame un autovalore alla volta.

   Per ogni i, $(i = 1, \dots, h)$, poniamo $\lambda_i = \lambda$ e procediamo come segue:

3.1) Sia $n = n - ma(\lambda)$. Poniamo:
   $r_0 = ma(\lambda)\,,$
   $r_1 = \text{car}(A - \lambda I) - \tilde{n}\,,$
   $r_2 = \text{car}((A - \lambda I)^2) - \tilde{n}\,,$
   $\dots$
   $r_{\nu-1} = \text{car}((A - \lambda I)^{\nu-1}) - \tilde{n} \neq 0\,,$
   $r_\nu = \text{car}((A - \lambda I)^{\nu}) - \tilde{n} = 0\,.$
   Come è evidente, il procedimento si arresta al primo indice ν per il quale risulti $r_\nu = 0$.

3.2) Poniamo:
   $b_j = r_{j-1} - r_j$ per $j = 1, \dots, \nu$
   $b_{\nu+1} = 0\,.$

3.3) Per $k = 1, \dots, \nu$ poniamo:
   $d_k = b_k - b_{k+1}\,.$

3.4) d_k è il numero di blocchi di Jordan di ordine k che compaiono sulla diagonale di $J(\lambda)$.
   $J(\lambda)$ è la matrice diagonale a blocchi che ha sulla diagonale:
   d_ν blocchi $J_\nu(\lambda)\,,$
   $d_{\nu-1}$ blocchi $J_{\nu-1}(\lambda)\,,$
   $\dots$
   d_2 blocchi $J_2(\lambda)\,,$
   d blocchi $J_1(\lambda) = (\lambda)\,.$

4) Iterare il punto 3) per tutti gli autovalori di A.

5) Costruzione di J_A come matrice diagonale a blocchi, avente sulla diagonale le matrici $J(\lambda_1), \dots, J(\lambda_h)$.

Esempio 7.4.28 *Determinare la forma di Jordan della matrice*

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ -1 & 1 & -1 & 1 & 0 \\ 1 & 1 & 3 & 0 & 0 \\ 1 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

1) $A - tI = \begin{pmatrix} 2-t & 0 & 0 & 0 & 0 \\ -1 & 1-t & -1 & 1 & 0 \\ 1 & 1 & 3-t & 0 & 0 \\ 1 & 0 & 0 & 2-t & 0 \\ 0 & 0 & 0 & 0 & -1-t \end{pmatrix}$

$p_A(t) = \det(A - tI) = (-1-t)(2-t)^4$.

2) *Dunque gli autovalori sono* -1 *e* 2; $ma(-1) = 1$, $ma(2) = 4$.

3) **Autovalore $\lambda = -1$.**

3.1) $\tilde{n} = 5 - ma(-1) = 4$. *Poniamo:*
$r_0 = 1$.

$$r_1 = \text{car}(A + I) - n = \text{car}\begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 1 & 0 \\ 1 & 1 & 4 & 0 & 0 \\ 1 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} - 4 = 4 - 4 = 0.$$

3.2) $\nu = 1$. *Per* $j = 1, \ldots, \nu$ *poniamo:*
$b_1 = r_0 - r_1 = 1$,
$b_2 = 0$.

3.3) $d_1 = b_1 - b_2 = 1$.
Quindi $J(-1)$ *ha ordine* 1 *ed è fatta da un solo blocco di ordine* 1.
$J(-1) = (-1)$.

Autovalore $\lambda = 2$.

3.1) *Sia* $n = n - ma(2) = 5 - 4 = 1$.
$r_0 = ma(2) = 4$.

$$r_1 = \text{car}(A - 2I) - n = \text{car}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -3 \end{pmatrix} - 1 = 4 - 1 = 3.$$

$$r_2 = \text{car}((A - 2I)^2) - n = \text{car}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 9 \end{pmatrix} - 1 = 3 - 1 = 2.$$

$$r_3 = \text{car}((A-2I)^3) - n = \text{car}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 27 \end{pmatrix} - 1 = 2 - 1 = 1.$$

$$r_4 = \text{car}((A-2I)^4) - n = \text{car}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 81 \end{pmatrix} - 1 = 1 - 1 = 0.$$

3.2) Per $j = 1, \dots, 4$ *poniamo:*
$b_4 = r_3 - r_4 = 1\,,$
$b_3 = r_2 - r_3 = 1\,,$
$b_2 = r_1 - r_2 = 1\,,$
$b_1 = r_0 - r_1 = 1\,,$
$b_5 = 0\,.$

3.3) Per $k = 1, \dots, 4$ *poniamo:*
$d_4 = b_4 - b_5 = 1,$
$d_3 = b_3 - b_4 = d_2 = b_2 - b_3 = d_1 = b_1 - b_2 = 0.$
Quindi $J(2)$ *è formata da un blocco di ordine 4,* $J(2) = J_4(2)$.

5) $$J_A = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix} - 1 = 2 - 1 = 1.$$

Ricetta per trovare una base di Jordan di una matrice. Sia A una matrice quadrata di ordine n.

1) Determinare il polinomio caratteristico $p_A(t)$ di A.

2) Determinare tutte le radici di $p_A(t)$, cioè gli autovalori di A, con la loro molteplicità algebrica. Siano esse $\lambda_1, \dots, \lambda_h$ e $ma(\lambda_i)$ sia la molteplicità algebrica di λ_i.

3) Si prende in esame un autovalore alla volta.

   Per ogni i, $(i = 1, \dots, h)$, poniamo $\lambda_i = \lambda$ e procediamo come segue:

3.1) Sia $\tilde{n} = n - ma(\lambda)$. Poniamo:
$r_0 = ma(\lambda)\,,$
$r_1 = \mathrm{car}(A - \lambda I) - \tilde{n}\,,$
$r_2 = \mathrm{car}((A - \lambda I)^2) - \tilde{n}\,,$
$\dots$
$r_{\nu-1} = \mathrm{car}((A - \lambda I)^{\nu-1}) - \tilde{n} \neq 0\,,$
$r_\nu = \mathrm{car}((A - \lambda I)^{\nu}) - \tilde{n} = 0\,.$
Come è evidente, il procedimento si arresta al primo indice ν per il quale risulti $r_\nu = 0$.

3.2) Poniamo $\delta_i = n - \mathrm{car}((A - \lambda I)^i) = \dim \mathrm{Ker}((A - \lambda I)^i)$.

3.3) Si determini una base $\mathcal{B}_{(\nu-1)}$ di $ker(T^{\nu-1})$, dove $T = A - \lambda I$.
Siccome $ker(T^{\nu-1}) \subseteq \mathrm{Ker}(T^\nu)$, sia

$$^{(\nu)}x_1, \dots, ^{(\nu)}x_{k(\nu)}$$

un completamento di $\mathcal{B}^{(\nu-1)}$ ad una base di $\mathrm{Ker}(T^\nu)$.
Si determini una base $\mathcal{B}^{(\nu-2)}$ di $ker(T^{\nu-2})$. Sia

$$^{(\nu-1)}x_1, \dots, ^{(\nu-1)}x_{k(\nu-1)}$$

un completamento di $\mathcal{B}^{(\nu-2)} \cup \{T(^{(\nu)}x_1), \dots, T(^{(\nu)}x_{k(\nu)})\}$.
Si determini una base $\mathcal{B}^{(\nu-3)}$ di $\mathrm{Ker}(T^{\nu-3})$. Sia

$$^{(\nu-2)}x_1, \dots, ^{(\nu-2)}x_{k(\nu-2)},$$

un completamento ad una base di $\mathrm{Ker}(T^{\nu-2})$ di

$$\mathcal{B}^{(\nu-3)} \cup \{T^2(^{(\nu)}x_1), \dots, T^2(^{(\nu)}x_{k(\nu)})\} \cup \{T(^{(\nu-1)}x_1), \dots, T(^{(\nu-2)}x_{k(\nu-1)})\}\,.$$

Al passo j si determina una base $\mathcal{B}^{(\nu-j)}$ di $\mathrm{Ker}(T^{\nu-j})$ e si considera $\mathcal{B}^{(\nu-j)} \cup \mathcal{T}_\lambda(j)$, dove

$$\mathcal{T}_\lambda(j) = \{y \in V \mid y = T^{j-1-l}(^{(\nu-l)}x_i)\,,\ 0 \le l \le j-1,\ 1 \le i \le k(\nu - l)\}\,.$$

Infine sia

$$^{(\nu-j+1)}x_1, \dots, ^{(\nu-j+1)}x_{k(\nu-j+1)}$$

un completamento di $\mathcal{B}^{(\nu-j)} \cup \mathcal{T}_\lambda(j)$ ad una base di $\mathrm{Ker}(T^{\nu-j-1})$.

Il processo si arresta a $j = \nu - 1$.

4) Ciascun $\mathcal{T}^\lambda(\nu-1)$ è una base di Jordan per l'autospazio generalizzato $W(\lambda)$. Poiché ν dipende da λ, sia $\nu(i)$ il valore di ν relativo a λ_i. Allora: $\mathcal{T}_{\lambda_1}(\nu(1)-1)\cup\cdots\cup\mathcal{T}_{\lambda_h}(\nu(h)-1)$ è una base di Jordan di tutto V.

7.5 Alcune applicazioni della forma di Jordan

Data una matrice $A \in \mathbf{M}_n(\mathbb{R})$, calcolarne le potenze A^m, specie per m abbastanza grande, è un procedimento che richiede molti calcoli.

Osservazione 7.5.1 *Se A è simile a B, anche A^m è simile a B^m.*

Infatti, se per ipotesi, $B = M^{-1}AM$, allora: $B^2 = M^{-1}AMM^{-1}AM = M^{-1}A^2M$. In generale, se $B^{m-1} = M^{-1}A^{m-1}M$, allora

$$B^m = B^{m-1}B = M^{-1}A^{m-1}MM^{-1}AM = M^{-1}A^mM\,.$$

Osservazione 7.5.2 *Se A è una matrice diagonale a blocchi:*

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & A_h \end{pmatrix},$$

è facile convincersi che:

$$A^m = \begin{pmatrix} A_1^m & 0 & \cdots & 0 \\ 0 & A_2^m & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & A_h^m \end{pmatrix},$$

cioè che la potenza di A è ancora una matrice diagonale a blocchi con i blocchi di A^m che sono le potenze dei blocchi di A.

È inoltre abbastanza facile calcolare la potenza di un blocco di Jordan

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \lambda & 1 \\ 0 & \cdots & \cdots & \cdots & \lambda \end{pmatrix}.$$

$$J_k(\lambda)^n = \begin{pmatrix} \lambda^n & 1 & 0 & \cdots & \binom{n}{k-1}\lambda^{n-k+1} \\ 0 & \lambda^n & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \lambda^n & n\lambda^{n-1} \\ 0 & \cdots & \cdots & 0 & \lambda^n \end{pmatrix},$$

dove $\binom{n}{k} = \frac{n!}{h!(n-h)!}$ *se* $h \leq n$, *mentre* $\binom{n}{k} = 0$ *se* $h > n$, *con* n, h *interi positivi.*

In altre parole, $J^k(\lambda)^n$ *ha su ciascuna riga lo sviluppo di* $(\lambda+1)^n$ *con il primo termine* λ_n *posto sulla diagonale principale. Se tale sviluppo ha più termini di quanti posti restano sulla riga, allora viene troncato; se invece restano posti liberi, ad essi viene assegnato valore 0.*

Esempio 7.5.3 $J_3(\lambda)^4 = \begin{pmatrix} \lambda^4 & 4\lambda^3 & 6\lambda^2 \\ 0 & \lambda^4 & 4\lambda^3 \\ 0 & 0 & \lambda^4 \end{pmatrix}$.

$$J_5(2)^3 = \begin{pmatrix} 2^3 & 3\cdot 2^2 & 3\cdot 2 & 1 & 0 \\ 0 & 2^3 & 3\cdot 2^2 & 3\cdot 2 & 1 \\ 0 & 0 & 2^3 & 3\cdot 2^2 & 3\cdot 2 \\ 0 & 0 & 0 & 2^3 & 3\cdot 2^2 \\ 0 & 0 & 0 & 0 & 2^3 \end{pmatrix} = \begin{pmatrix} 8 & 12 & 6 & 1 & 0 \\ 0 & 8 & 12 & 6 & 1 \\ 0 & 0 & 8 & 12 & 6 \\ 0 & 0 & 0 & 8 & 12 \\ 0 & 0 & 0 & 0 & 8 \end{pmatrix}.$$

Osservazione 7.5.4 $J_k(0)^k = 0$.

Sia $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ un polinomio. Con $p(A)$ si indica la matrice che si ottiene sostituendo A ad x:

$$p(A) = a_n A^n + a_{n-1}A^{n-1} + \cdot + a_1 A + a_0 I\,,$$

avendo posto $A^0 = I_n$.

Proposizione 7.5.5 *Se* A *è simile a* B*, allora* $p(A)$ *è simile a* $p(B)$ *per ogni polinomio* p.

Dim. Sia $B = M^{-1}AM$. Allora, per l'Osservazione 7.5.1, $B^m = M^{-1}A^m M$.

$$\begin{aligned} p(B) &= a_n B^n + a_{n-1}B^{n-1} + \cdots + a_1 B + a_0 I \\ &= a_n M^{-1}A^n M + a_{n-1}M^{-1}A^{n-1}M + \cdots + a_1 M^{-1}AM + a_0 M^{-1}M \\ &= M^{-1}(a_n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I)M = M^{-1}p(A)M\,. \end{aligned}$$

□

Osservazione 7.5.6 *Una diretta conseguenza della Osservazione 7.5.2 è che se A è una matrice diagonale a blocchi*

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & A_h \end{pmatrix},$$

allora

$$p(A) = \begin{pmatrix} p(A_1) & 0 & \cdots & 0 \\ 0 & p(A_2) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & p(A_h) \end{pmatrix}.$$

Possiamo a questo punto facilmente dimostrare il seguente teorema:

Teorema 7.5.7 (Hamilton-Cayley) *Sia $p_A(t)$ il polinomio caratteristico di una matrice $A \in \mathbf{M}_n(\mathbb{R})$. Allora:*

$$p_A(A) = 0\,.$$

Dim. Basta dimostrare che $p_A(J_A) = 0$ dove J_A è la forma di Jordan di A. Infatti sappiamo che A e J_A sono simili e quindi anche $p_A(A)$ è simile a $p_A(J_A)$. Se questa matrice è 0, anche $p_A(A) = 0$.

Ricordiamo (7.13) che:

$$J_A = \begin{pmatrix} J(\lambda_1) & 0 & \cdots & 0 \\ 0 & J(\lambda_2) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & J(\lambda_h) \end{pmatrix}.$$

Per l'Osservazione 7.5.6, si ha che:

$$p_A(J_A) = \begin{pmatrix} p(J(\lambda_1)) & 0 & \cdots & 0 \\ 0 & p(J(\lambda_2)) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & p(J(\lambda_h)) \end{pmatrix}.$$

Per $\lambda = \lambda_1, ..., \lambda_h$, si ha:

$$J(\lambda) = \begin{pmatrix} J_{k_1}(\lambda) & 0 & \cdots & 0 \\ 0 & J_{k_2}(\lambda) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & J_{k_m}(\lambda) \end{pmatrix},$$

$$p_A(J(\lambda)) = \begin{pmatrix} p(J_{k_1}(\lambda)) & 0 & \cdots & 0 \\ 0 & p(J_{k_2}(\lambda)) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & p(J_{k_m}(\lambda)) \end{pmatrix}.$$

Prendiamo in esame il generico $p_A(J_k(\lambda))$.
Ricordiamo che $p_A(t) = (t - \lambda)^{m_a(\lambda)} q(t)$. Infatti λ è radice di $p_A(t)$ con molteplicità $m_a(\lambda)$.

$$p_A(J_k(\lambda)) = (J_k(\lambda) - \lambda I)^{m_a(\lambda)} q(J_k(\lambda)).$$

Ma è facile verificare che $J_k(\lambda) - \lambda I = J_k(0)$.

Inoltre la molteplicità algebrica di λ, $m_a(\lambda)$, è uguale alla dimensione dell'autospazio generalizzato relativi a λ, dunque uguale all'ordine di $J(\lambda)$. Quindi $m_a(\lambda) \geq k$, dunque, per l'Osservazione 7.5.4, si ha che:

$$(J_k(0))^{m_a(\lambda)} = 0.$$

Ciò implica che:

$$\begin{aligned} p_A(J_k(\lambda)) &= (J_k(\lambda) - \lambda I)^{m_a(\lambda)} q(J_k(\lambda)) \\ &= (J_k(0))^{m_a(\lambda)} q(J_k(\lambda)) = 0. \end{aligned}$$

Poiché questo vale per ogni autovalore λ, il teorema è dimostrato. □

7.6 Forma di Jordan reale

Supponiamo che A sia una matrice di $\mathbf{M}_n(\mathbb{R})$, ma i suoi valori non siano tutti reali. Supponiamo che vi sia $\lambda \in \mathbb{C}$, non reale, tra le radici del polinomio caratteristico di A. Poiché però questo è a coefficienti reali, anche $\overline{\lambda}$ è autovalore (cfr. Proposizione 2.3.12).

Enunciamo quanto segue senza dimostrazione.

Proposizione 7.6.1 *I blocchi di Jordan $J_k(\lambda)$ in J_A sono in corrispondenza biunivoca con i blocchi $J_k(\overline{\lambda})$.*

Sia $\lambda = a + ib$. Siano

$$\omega(\lambda) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}, \quad \mathbf{1} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Proposizione 7.6.2 *La matrice*

$$\begin{pmatrix} J_k(\lambda) & 0 \\ 0 & J_k(\overline{\lambda}) \end{pmatrix}$$

di ordine $2k$ è simile a:

$$\Omega_k(\lambda) = \begin{pmatrix} \omega(\lambda) & \mathbf{1} & 0 & \cdots & 0 \\ 0 & \omega(\lambda) & \mathbf{1} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \omega(\lambda) & \mathbf{1} \\ 0 & \cdots & \cdots & \cdots & \omega(\lambda) \end{pmatrix}$$

che ha sulla diagonale k blocchi

$$\omega(\lambda) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

La forma di Jordan reale allora si trova ricavando la forma complessa, permutando i blocchi in modo che risultino adiacenti blocchi di Jordan dello stesso ordine relativi ad autovalori coniugati e sostituendo la coppia di tali blocchi, sulla diagonale, con una matrice reale del tipo $\Omega_k(\lambda)$ di ordine $2k$.

Esempio 7.6.3 $A = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$.
Il polinomio caratteristico è $t^2 + t + 1$.
Gli autovalori sono $\lambda = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$ e $\overline{\lambda} = -\frac{1}{2} - i\frac{\sqrt{3}}{2}$.
Abbiamo $J(A) = \begin{pmatrix} \lambda & 0 \\ 0 & \overline{\lambda} \end{pmatrix}$.
La forma di Jordan reale è

$$\begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

In questo caso non è un gran vantaggio.

7.7 Alcune dimostrazioni

Dimostriamo a questo punto l'Osservazione 7.4.15 utilizzate precedentemente. Richiamiamole.

1) $T^k_{\lambda i}$ e $T^h_{\lambda j}$ commutano. Ricordiamo che due endomorfismi F e G commutano se $F \circ G = G \circ F$.

Dim. Basta far vedere che T_{λ_i} e T_{λ_j} commutano. Siccome $T_{\lambda i} = F - \lambda_i I, T_{\lambda j} = F - \lambda_j I$, abbiamo

$$\begin{aligned} T_{\lambda_i} T_{\lambda_j} &= (F - \lambda_i I)(F - \lambda_j I) = F^2 - (\lambda_i + \lambda_j)F + (\lambda_i \lambda_j I) \\ &= F^2 - (\lambda_j + \lambda_i)F + (\lambda_j \lambda_i I) = (F - \lambda_j I)(F - \lambda_i I) = T_{\lambda_j} T_{\lambda_i} \,. \end{aligned}$$

□

2) $F_{|W(\lambda)}$ è un endomorfismo di $W(\lambda)$ che ha come unico autovalore λ.

Dim. Supponiamo per assurdo che $\mu \neq \lambda$ sia autovalore di $F_{|W(\lambda)}$.
Sia x autovettore non nullo in $W(\lambda)$, cioè $T_\mu(x) = 0$. Poiché $W(\lambda) = \text{Ker}(T^k_\lambda)$, $T^k_\lambda(x) = 0$. Esiste un indice h tale che $T^h_\lambda(x) = y \neq 0$, mentre $T^{h+1}_\lambda(x) = T_\lambda(y) = 0$. Allora $y \neq 0$ è autovettore relativo a λ perché $T_\lambda(y) = 0$, ma anche autovettore non nullo relativo a μ. Infatti

$$T_\mu(y) = T_\mu(T^h_\lambda(x)) = T^h_\lambda(T_\mu(x)) = 0$$

poiché T^k_λ e T_μ commutano. Ma ciò è assurdo per il Corollario 7.2.16. □

3) Se $i \neq j$, $W(\lambda_i) \cap W(\lambda_j) = \{0\}$.

Dim. Si procede in maniera simile alla precedente.
Sia $0 \neq x \in W(\lambda_i) \cap W(\lambda_j)$.
Per le definizioni, esistono interi h e k tali che: $T^k_{\lambda_i}(x) = 0 = T^h_{\lambda_j}(x)$.
Esiste un indice p tale che $y = T^p_{\lambda_i}(x) \neq 0$ ma $T^{p+1}_{\lambda_i}(x) = T_{\lambda_i}(y) = 0$.
Siccome $T^h_{\lambda_j}(x) = 0$, anche $T^h_{\lambda_j}(y) = T^h_{\lambda_j}(T^p_{\lambda_i}(x)) = T^p_{\lambda_i}(T^h_{\lambda_j}(x)) = 0$.
Esiste un indice q tale che $T^q_{\lambda_j}(y) \neq 0$ ma $T^{q+1}_{\lambda_j}(y) = 0$.
Sia $z = T^q_{\lambda_j}(y)$. $z \neq 0$, ma $T_{\lambda_j}(z) = T^{q+1}_{\lambda_j}(y) = 0$, quindi z è autovettore relativo a λ_j.

$$T_{\lambda_i}(z) = T_{\lambda_i}(T^q_{\lambda_j}(y)) = T^q_{\lambda_j}(T_{\lambda_i}(y)) = 0 \,,$$

perché $T_{\lambda_i}(y) = 0$. Quindi z è autovettore relativo a λ_i. Assurdo se $\lambda_i \neq \lambda_j$.

□

4) $V = W(\lambda_i) \oplus \mathrm{Im}T_{\lambda_i}^{\nu_i}$, dove ν_i è tale che $W(\lambda_i) = V_{\nu_i}(\lambda_i)$.

Dim. Basta dimostrare che $W(\lambda_i) \cap \mathrm{Im}T_{\lambda_i}^{\nu_i} = 0$.
Sia $0 \neq x \in W(\lambda_i) \cap \mathrm{Im}T_{\lambda_i}^{\nu_i}$. Poiché $x \in \mathrm{Im}T_{\lambda_i}^{\nu_i}$, sia y tale che: $x = T_{\lambda_i}^{\nu_i}(y)$. Poiché $x \in W(\lambda_i)$, $T_{\lambda_i}^{\nu_i}(x) = 0$. Ma allora $T_{\lambda_i}^{\nu_i}(T_{\lambda_i}^{\nu_i}(y)) = T_{\lambda_i}^{2\nu_i}(y) = 0$. Quindi $y \in \mathrm{Ker}T_{\lambda_i}^{2\nu_i}$. Ma $\mathrm{Ker}T_{\lambda_i}^{\mu} = \mathrm{Ker}T_{\lambda_i}^{\nu_i} = W(\lambda_i)$ per ogni $\mu \geq \nu_i$. Quindi $y \in \mathrm{Ker}T_{\lambda_i}^{\nu_i}$ e $x = T_{\lambda_i}^{\nu_i}(y) = 0$. □

5) $W(\lambda_i)$ e $\mathrm{Im}T_{\lambda_i}^{\nu_i}$ sono sottospazi F-invarianti e anche T_{λ_j}-invarianti per ogni autovalore λ_j di F.

Dim. È una semplice verifica tenuto conto del fatto che F, T_{λ_i} e T_{λ_j} commutano tra loro. □

6) Se $i \neq j$, $\mathrm{Ker}(T_{\lambda_j}^h) = V_h(\lambda_j) \subseteq \mathrm{Im}T_{\lambda_i}^{\nu_i}$.

Dim. Sia $y \neq 0$, $y \in \mathrm{Ker}(T_{\lambda_j}^h)$, ma $y \notin \mathrm{Im}T_{\lambda_i}^{\nu_i}$. Per la 4), $y = x + v$, con $x \in W(\lambda_i)$ e $v \in \mathrm{Im}T_{\lambda_i}^{\nu_i}$. $x \neq 0$. Poiché $T_{\lambda_j}^h(y) = 0$, per l'unicità della decomposizione e per la $T_{\lambda_j}^h$-invarianza di $W(\lambda_i)$ e di $\mathrm{Im}T_{\lambda_i}^{\nu_i}$, si ha che:

$$T_{\lambda_j}^h(x) = T_{\lambda_j}^h(v) = 0\,.$$

Allora $x \in \mathrm{Ker}(T_{\lambda_j}^h) \cap W(\lambda_i)$, quindi *a fortiori* $x \in W(\lambda_j) \cap W(\lambda_i)$. Ma, per la 3), se $i \neq j$, tale intersezione è 0 e pertanto si ha un assurdo. □

7.8 Esercizi

Diagonalizzazione di applicazioni lineari e matrici

1) Diagonalizzare, quando possibile, le seguenti matrici reali:

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & -4 & 8 \\ 7 & -7 & 18 \\ 2 & -2 & 5 \end{pmatrix}.$$

2) Stabilire per quali valori del parametri reali h e k le seguenti matrici sono diagonalizzabili:

$$A_h = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & h & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 1 & 0 & 2 \end{pmatrix}, \quad B_k = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & -1 & 0 & k \end{pmatrix}.$$

3) Sia $f : \mathbb{R}^3 \to \mathbb{R}^3$ l'applicazione lineare definita da:

$$\begin{aligned} f(e_1) &= 4e_1 + 5e_2 - 2e_3\,, \\ f(e_2) &= -2e_1 - 2e_2 + e_3\,, \\ f(e_3) &= -e_1 - e_2 + e_3\,. \end{aligned}$$

Stabilire se f è diagonalizzabile.

4) Data la matrice

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix},$$

sia $L_A : \mathbb{K}^3 \to \mathbb{K}^3$ l'applicazione lineare associata ad A. Si stabilisca se L_A è diagonalizzabile nei casi $\mathbb{K} = \mathbb{R}$ e $\mathbb{K} = C$.

5) Data la matrice a coefficienti complessi

$$A = \begin{pmatrix} 0 & -1 & -3 \\ -2 & 1 & 1 \\ 2 & 3 & 3 \end{pmatrix} \in M_3(\mathbb{C}),$$

determinare una base di $\mathbb{C}^3$ di autovettori per A e scrivere una matrice diagonale simile ad essa.

6) Sia f un endomorfismo di uno spazio vettoriale reale o complesso di dimensione finita. Stabilire se f è diagonalizzabile in ciascuno dei seguenti casi:

1. $f^2 = f$,
2. $f^2 = id$,
3. $f^2 = 0$,
4. $f^2 = -id$.

Capitolo 8

PRODOTTI SCALARI E GRUPPI ORTOGONALI

8.1 Forme bilineari e prodotti scalari

Definizione 8.1.1 *Siano V e W due spazi vettoriali reali. Una applicazione $b : V \times W \to \mathbb{R}$ è una* forma bilineare *se valgono le seguenti proprietà:*

i) $b(\lambda v_1 + \mu v_2, w) = \lambda b(v_1, w) + \mu b(v_2, w)$,

ii) $b(v, \lambda w_1 + \mu w_2) = \lambda b(v, w_1) + \mu b(v, w_2)$,

per ogni $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$, $\lambda, \mu \in \mathbb{R}$.

Esempi. 1) $b : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$ tale che $b(a, (b, c)) = a(b + c)$.
2) $b : \mathbb{R}[x] \times \mathbb{R}[x] \to \mathbb{R}$ tale che $b(p(x), q(x)) = p(1)q(0)$.

Osservazione 8.1.2 *Per ogni fissato $w \in W$, l'applicazione $b(\cdot, w) : V \to \mathbb{R}$ tale che $b(\cdot, w)(v) = b(v, w)$ è lineare. Lo stesso si ottiene se fisso $v \in V$.*
In particolare, $b(0, w) = b(v, 0) = 0$ per ogni $v \in V$ e $w \in W$.

Definizione 8.1.3 *Una forma bilineare $b : V \times V \to \mathbb{R}$ si dice* prodotto scalare *se:*

$$b(v_1, v_2) = b(v_2, v_1) \qquad \textit{per ogni } v_1, v_2 \in V\,.$$

Esempio 8.1.4 *1)* $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *tale che*

$$p\left(\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}\right) = (a_1, \ldots, a_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + \cdots + a_n b_n \,.$$

Questo viene detto prodotto scalare canonico *di* $\mathbb{R}^n$*.*
2) $s : \mathbb{R}[x] \times \mathbb{R}[x] \to \mathbb{R}$ *tale che*

$$s(p, q) = \int_0^1 p(x) q(x) dx \,.$$

Notazione. D'ora in avanti, se s è un prodotto scalare, scriveremo $\langle a, b \rangle$ per $s(a, b)$.

Definizione 8.1.5 *Un prodotto scalare* $\langle \, , \rangle$ *si dice* non degenere *se:*

$$\langle u, v \rangle = 0 \quad \forall v \in V \iff u = 0 \,.$$

In altre parole, se l'unico elemento di V *che moltiplicato scalarmente per tutti gli altri dà sempre come risultato 0 è lo zero di* V*.*

Definizione 8.1.6 *Un prodotto scalare* $\langle \, , \rangle$ *si dice* definito positivo *se si ha* $\langle v, v \rangle \geq 0$*, per ogni* $v \in V$*, e* $\langle v, v \rangle = 0$ *se e solo se* $v = 0$*. Questo equivale a:*

$$\langle v, v \rangle > 0 \,, \forall v \neq 0 \,.$$

Il numero $|v| = \sqrt{\langle v, v \rangle}$ *è un reale positivo e* $|v| = 0$ *se e solo se* $v = 0$*. Esso viene anche detta* norma *di* v*.*

Osservazione 8.1.7 *Se un prodotto scalare* $\langle \, , \rangle$ *è definito positivo, allora esso è necessariamente non degenere.*

Gli esempi in 8.1.4 sono prodotti scalari definiti positivi.

Vedremo più avanti un criterio per determinare se un dato prodotto scalare sia definito positivo.

8.2 Prodotti scalari e matrici

Sia V uno spazio vettoriale a dimensione finita (d'ora in avanti tutti gli spazi vettoriali che considereremo saranno a dimensione finita). Sia $\mathcal{B} = \{b_1, \ldots, b_n\}$ una base di V. Allora ad un prodotto scalare $\langle\, ,\rangle$ definito su V può essere associata una matrice A nel seguente modo:

$$A = (a_{ij}) \text{ con } a_{ij} = \langle b_i, b_j \rangle \,. \tag{8.1}$$

Osservazione 8.2.1 *Per la definizione di prodotto scalare,* $\langle b_i, b_j\rangle = \langle b_j, b_i\rangle$. *Quindi* $a_{ij} = a_{ji}$ *e* A *è una matrice simmetrica.*

Esempio 8.2.2 *1)* $V = \mathbb{R}^n$. $\mathcal{C} = \{e_1, \ldots, e_n\}$ *la base canonica. Consideriamo il prodotto scalare canonico di* $\mathbb{R}^n$.

$$\langle e_i, e_j\rangle = \delta_{ij} \,,$$

dove $\delta_{ij} = 1$ *se* $i = j$ *e* $\delta_{ij} = 0$ *se* $i \neq j$.

Quindi la matrice associata a $\langle\, ,\rangle$ *nella base canonica è* I_n.

2) $V = \mathbb{R}_2[x]$, *lo spazio dei polinomi a coefficienti reali di grado minore o uguale a 2. Come prodotto scalare consideriamo quello dell'Esempio 8.1.4, 2). Su* V *consideriamo la base* $\mathcal{B} = \{1, x, x_2\}$.

Scriviamo la matrice A *associata a* $\langle\, ,\rangle$ *nella base* $\mathcal{B}$.

$$\begin{array}{ll} a_{11} = \langle 1, 1\rangle = \int_0^1 1dx = 1\,, & a_{12} = \langle 1, x\rangle = \int_0^1 xdx = \frac{1}{2}\,, \\ a_{13} = \langle 1, x^2\rangle = \int_0^1 x^2dx = \frac{1}{3}\,, & a_{22} = \langle x, x\rangle = \int_0^1 x^2dx = \frac{1}{3}\,, \\ a_{23} = \langle x, x^2\rangle = \int_0^1 x^3dx = \frac{1}{4}\,, & a_{33} = \langle x^2, x^2\rangle = \int_0^1 x^4dx = \frac{1}{5}\,. \end{array}$$

Poiché A *è simmetrica si ha che:*

$$A = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix} \,.$$

Ricordiamo che, fissata una base $\mathcal{B}$ di V, ad ogni vettore $v \in V$ possiamo far corrispondere la n-upla delle sue coordinate.

Allora, se v ha coordinate $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, e w ha coordinate $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, è facile verificare che:

$$\langle v, w\rangle = (x_1, \dots, x_n)A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} ,$$

dove A è la matrice associata a $\langle\, ,\rangle$ nella base fissata $\mathcal{B}$.

Vogliamo ora studiare come la matrice associata ad un prodotto scalare dipenda dalla base scelta.

Sia $\mathcal{B}'$ un'altra base di V e sia M la matrice di passaggio dalle coordinate in base $\mathcal{B}'$ a quelle in base B. Cioè:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = M \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} ,$$

dove $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ sono le coordinate, appunto, nella base $\mathcal{B}'$.

Se indichiamo con A' la matrice associata a $\langle\, ,\rangle$ nella base $\mathcal{B}'$, si ha:

$$\langle v, w\rangle = (x_1, \dots, x_n)A \begin{pmatrix} y'_1 \\ \vdots \\ y'_n \end{pmatrix} = (x'_1, \dots, x'_n)\, {}^tMAM \begin{pmatrix} y'_1 \\ \vdots \\ y'_n \end{pmatrix} .$$

Ma deve risultare anche che:

$$\langle v, w\rangle = (x'_1, \dots, x'_n)A' \begin{pmatrix} y'_1 \\ \vdots \\ y'_n \end{pmatrix} .$$

Abbiamo così provato la seguente proposizione.

Proposizione 8.2.3 *Sia V uno spazio vettoriale su cui è dato un prodotto scalare $\langle\, ,\rangle$. Sia $\mathcal{B}$ una sua base e A la matrice associata a $\langle\, ,\rangle$ nella base $\mathcal{B}$. Sia $\mathcal{B}'$ un'altra base di V. Se A' indica la matrice associata a $\langle\, ,\rangle$ nella base $\mathcal{B}'$, allora si ha:*

$$A' = {}^tMAM\,, \tag{8.2}$$

dove M è la matrice di passaggio dalle coordinate nella base $\mathcal{B}'$ in quelle nella base $\mathcal{B}$.

Osserviamo che A' in effetti è simmetrica. Ma attenzione: A e A' *non* sono simili. In generale, infatti, ${}^tM$ è ben diversa da M^{-1}.

Definizione 8.2.4 *Due matrici quadrate A, B si dicono* congruenti *se esiste una matrice M* invertibile *tale che:*

$$B = {}^tMAM\,.$$

Proposizione 8.2.5 *Sia A la matrice associata al prodotto scalare $\langle\, ,\rangle$ su V in una certa base $\mathcal{B}$. Allora $\langle\, ,\rangle$ è non degenere se e solo se* $\det A \neq 0$.

Dim. Per il teorema di Cramer, si ha che $\det A = 0$ se e solo se esiste una n-upla $(a_1, \dots, a_n)$ non nulla tale che

$$A \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Sia v_0 il vettore di V che ha, in base $\mathcal{B}$, come coordinate $(a_1, \dots, a_n)$. Ovviamente $v_0 \neq 0$. Ma, per ogni $v \in V$ si ha:

$$\langle v, v_0\rangle = (x_1, \dots, x_n) A \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = (x_1, \dots, x_n) \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = 0\,,$$

dove $(x_1, \dots, x_n)$ sono le coordinate di v in base $\mathcal{B}$.

Quindi, se $\det A = 0$, allora $\langle\, ,\rangle$ è degenere.

Viceversa, se $\langle\, ,\rangle$ fosse degenere, esisterebbe un vettore di V non nullo v_0 tale che, per ogni $v \in V$ si ha:

$$\langle v, v_0\rangle = (x_1, \dots, x_n) A \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = 0\,.$$

Se anche un solo elemento della n-upla $\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = A \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ fosse diverso da zero, ad esempio c_1, allora $\langle b_1, v_0 \rangle = c_1 \neq 0$, dove b_1 è il primo elemento della base $\mathcal{B}$. Ma poiché per ipotesi, il risultato deve essere zero, allora, necessariamente si deve avere:

$$A \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} .$$

Quindi, per il teorema di Cramer, $\det A = 0$. □

Definizione 8.2.6 *Il sottospazio V_0 di V*

$$V_0 = \{v \in V \mid \langle v, u \rangle = 0 \text{ per ogni } u \in V\}$$

si dice nucleo *di $\langle \, , \rangle$.*

Teorema 8.2.7 (Riesz) *Sia $\langle \, , \rangle$ un prodotto scalare non degenere su V e sia $f : V \to \mathbb{R}$ una applicazione lineare. Allora esiste un vettore $v_0 \in V$ tale che $f(v) = \langle v_0, v \rangle$ per ogni $v \in V$.*

Dim. Fissata una base $\mathcal{B}$ di V, sia M la matrice associata alla applicazione lineare f in tale base $\mathcal{B}$. M è di tipo $1 \times n$, $M = (m_1, \dots, m_n)$.

Sia A la matrice associata a $\langle \, , \rangle$ nella base $\mathcal{B}$. Sappiamo che $\det A \neq 0$ perché $\langle \, , \rangle$ è non degenere. Sia v_0 il vettore che, nella base $\mathcal{B}$, ha coordinate

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = A^{-1} \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} = A^{-1} \cdot {}^tM .$$

Quindi si ha che ${}^tM = A \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ e dunque, poiché A è simmetrica,

$$M = (a_1, \dots, a_n)\, {}^tA = (a_1, \dots, a_n) A .$$

Allora, per ogni $v \in V$ si ha:

$$f(v) = (m_1, \dots, m_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = M \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (a_1, \dots, a_n) A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \langle v_0, v \rangle \,,$$

dove $(x_1, \dots, x_n)$ sono le coordinate di v nella base $\mathcal{B}$. □

Teorema 8.2.8 (di rappresentazione) *Sia $\langle\,,\rangle$ un prodotto scalare non degenere su V e sia $(\,,)$ un qualsiasi altro prodotto scalare. Allora esiste un unico endomorfismo $f : V \to V$ tale che:*

$$(u, v) = \langle f(u), v \rangle$$

per ogni $u, v \in V$.

Dim. Fissata una base $\mathcal{B}$ di V, siano N e A le matrici associate in tale base rispettivamente a $(\,,)$ e a $\langle\,,\rangle$. Sia f l'endomorfismo di V espresso, nella base $\mathcal{B}$, dalla matrice $A^{-1}N$.

In coordinate, si ha:

$$f(u) = A^{-1} N \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \,,$$

$$\langle f(u), v \rangle = (x_1, \dots, x_n)\, {}^t(A^{-1}N) A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \,.$$

Sappiamo che:

$${}^t(A^{-1}N) = {}^tN\, {}^t(A^{-1}) = {}^tN({}^tA)^{-1} = NA^{-1} \,,$$

avendo utilizzato che le matrici A e N simmetriche. In conclusione si ha:

$$\begin{aligned} \langle f(u), v \rangle &= (x_1, \dots, x_n)\, {}^t(A^{-1}N) A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = (x_1, \dots, x_n) N A^{-1} A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ &= (x_1, \dots, x_n) N \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = (u, v) \,. \end{aligned}$$

□

8.3 Basi ortogonali e basi ortonormali

Definizione 8.3.1 *Sia* $\langle\, ,\rangle$ *un prodotto scalare su* V. *Due vettori di* V, u *e* v *si dicono* ortogonali *se* $\langle u, v\rangle = 0$.

Definizione 8.3.2 *Se* W *è un sottoinsieme di* V *(non necessariamente un sottospazio) l'insieme* $W^\perp$ *così definito:*

$$W^\perp = \{v \in V \mid \langle v, w\rangle = 0, \forall w \in W\}$$

si dice sottospazio ortogonale *di* W.

Osservazione 8.3.3 *Qualunque sia* W, $W^\perp$ *è sempre un sottospazio vettoriale di* V. *Infatti, se* $v_1, v_2 \in W^\perp$, *allora per ogni* $\lambda, \mu \in \mathbb{R}$, *si ha:*

$$\langle \lambda v_1 + \mu v_2, w\rangle = \lambda\langle v_1, w\rangle + \mu\langle v_2, w\rangle = 0\,, \quad \forall w \in W\,.$$

Quindi anche $\lambda v_1 + \mu v_2 \in W^\perp$.

Definizione 8.3.4 *Sia* $\langle\, ,\rangle$ *come al solito un prodotto scalare su* V. *Una base* $\{b_1, \ldots, b_n\}$ *di* V *si dice* base ortogonale *se* $\langle b_i, b_j\rangle = 0$ *per ogni* $i \neq j$.

Osservazione 8.3.5 *La matrice associata a* $\langle\, ,\rangle$ *in una base ortogonale è una matrice diagonale.*

Definizione 8.3.6 *Sia* $\langle\, ,\rangle$ *un prodotto scalare su* V. *Una base* $\{b_1, \ldots, b_n\}$ *di* V *si dice* base ortonormale *se* $\langle b_i, b_j\rangle = 0$ *se* $i \neq j$ *e* $\langle b_i, b_i\rangle = 1$ *per ogni* i.

Osservazione 8.3.7 *La matrice associata a* $\langle\, ,\rangle$ *in una base ortonormale è la matrice identità.*

Teorema 8.3.8 *Dato un prodotto scalare* $\langle\, ,\rangle$ *sullo spazio vettoriale* V, *esiste sempre una base di* V *ortogonale.*

Dim. Per induzione su n, la dimensione dello spazio vettoriale V.

Se $n = 1$, allora ogni elemento non nullo di V è una base ortogonale.

Supponiamo che il teorema valga per ogni spazio vettoriale di dimensione $n - 1$; dimostriamo che vale per ogni spazio di dimensione n.

Distinguiamo due casi:

1) $\langle v, v\rangle = 0$ per ogni $v \in V$. In questo caso ogni base è ortogonale.
2) Esiste $v_1 \in V$ tale che $\langle v_1, v_1\rangle \neq 0$.

Sia allora V_1 il sottospazio di V generato da v_1; si ha che:

$$V = V_1 \oplus V_1^{\perp}\,.$$

Infatti, per prima cosa dimostriamo che $V = V_1 + V_1^{\perp}$, cioè che ogni elemento v di V si può scrivere come somma di un multiplo di v_1 con un elemento di $V_1^{\perp}$. Sia $v \in V$. Poniamo: $u = \frac{\langle v, v_1\rangle}{\langle v_1, v_1\rangle} v_1$. Poiché u è un multiplo di v_1, $u \in V_1$.

Sia $w = v - u$. Allora:

$$\begin{aligned}\langle w, v_1\rangle &= \langle v - u, v_1\rangle = \langle v, v_1\rangle - \langle u, v_1\rangle \\ &= \langle v, v_1\rangle - \frac{\langle v, v_1\rangle}{\langle v_1, v_1\rangle}\langle v_1, v_1\rangle = 0\,.\end{aligned}$$

Quindi $w \in V_1^{\perp}$.

Resta da far vedere che $V_1 \cap V_1^{\perp} = \{0\}$.

Sia $v = \lambda v_1 \in V_1 \cap V_1^{\perp}$. Poiché $\lambda v_1 \in V_1^{\perp}$, deve risultare $\langle \lambda v_1, v_1\rangle = 0$, per la definizione di $V_1^{\perp}$. Ma $\langle \lambda v_1, v_1\rangle = \lambda\langle v_1, v_1\rangle = 0$ se e solo se $\lambda = 0$ (ricordiamo che $\langle v_1, v_1\rangle \neq 0$).

Quindi $v = 0$.

Allora $\dim V_1^{\perp} = n - 1$. Per ipotesi induttiva sia $\mathcal{B} = \{b_1, \ldots, b_{n-1}\}$ una base ortogonale di $V_1^{\perp}$. Allora $\mathcal{B}' = \{b_1, \ldots, b_{n-1}, v_1\}$ è una base di V ortogonale. □

Per avere basi ortonormali abbiamo bisogno di un requisito in più per il prodotto scalare $\langle\, ,\rangle$. Vale infatti il seguente risultato

Teorema 8.3.9 *Sia $\langle\, ,\rangle$ un prodotto scalare su V. Esiste una base ortonormale se e solo se $\langle\, ,\rangle$ è definito positivo.*

Dim. Se il prodotto scalare $\langle\, ,\rangle$ ammette una base ortonormale, esso è chiaramente definito positivo. Infatti, sia $\mathcal{B}$ una tale base. La matrice associata a $\langle\, ,\rangle$ in questa base è la matrice identità.

Se $(x_1, \ldots, x_n)$ sono le coordinate di un vettore v nella base $\mathcal{B}$, si ha:

$$\langle v, v\rangle = x_1^2 + \cdots + x_n^2\,.$$

Quindi $\langle v, v\rangle \geq 0$; inoltre $\langle v, v\rangle = 0$ se e solo se $x_i = 0$, $i = 1, \ldots, n$ e cioè se $v = 0$.

Viceversa, supponiamo che $\langle\, ,\rangle$ sia definito positivo e sia $\mathcal{B} = \{b_1, \ldots, b_n\}$ una base di V. Se $v \in V$, possiamo definire la sua norma $|v| = \sqrt{\langle v, v\rangle}$. Ricordiamo che $|v|$ è un numero reale positivo e $|v| = 0$ se e solo se $v = 0$. Dunque, per ogni $i = 1, \ldots, n$, $|b_i| \neq 0$ e i vettori $b_i' = \frac{b_i}{|b_i|}$ hanno norma 1. Ne segue che $\mathcal{B}' = \{b_1', \ldots, b_n'\}$ è una base ortonormale di V. □

8.3.1 Ortogonalizzazione di Gram-Schmidt

Data una base $\mathcal{B} = \{b_1, \dots, b_n\}$ di uno spazio vettoriale V (in generale di un sottospazio vettoriale V di uno spazio U) su cui è definito un prodotto scalare definito positivo, si può costruire una base ortonormale di V con il seguente metodo, detto anche processo di *ortogonalizzazione di Gram-Schmidt*.

Poniamo:

$$\begin{aligned}
&b'_1 = \tfrac{b_1}{|b_1|}\,, && |b'_1| = 1 \text{ (ovviamente)}\,,\\
&b''_2 = b_2 - \langle b_2, b'_1\rangle b'_1\,, && b'_2 = \tfrac{b''_2}{|b''_2|}\,,\\
&b''_3 = b_3 - \langle b_3, b'_1\rangle b'_1 - \langle b_3, b'_2\rangle b'_2\,, && b'_3 = \tfrac{b''_3}{|b''_3|}\,,\\
&\dots && \dots\\
&b''_i = b_i - \langle b_i, b'_1\rangle b'_1 - \dots - \langle b_i, b'_{i-1}\rangle b'_{i-1}\,, && b'_i = \tfrac{b''_i}{|b''_i|}\,.
\end{aligned}$$

Ad ogni passo, si sottraggono all'i-esimo elemento della base $\mathcal{B}$ i vettori b'_j (con $j < i$), precedentemente ottenuti, moltiplicati per $\langle b_i, b'_j\rangle$, che è un numero reale; poi il risultato di questa sottrazione viene diviso per la sua norma.

Dimostriamo ora che l'insieme $\mathcal{B}' = \{b'_1, \dots, b'_n\}$ è effettivamente una base ortonormale.

È un insieme di generatori perché ogni elemento di $\mathcal{B}$ si può scrivere come combinazione lineare dei vettori di $\mathcal{B}'$; ad esempio b_i:

$$b_i = \langle b_i, b'_1\rangle b'_1 + \dots + \langle b_i, b'_{i-1}\rangle b'_{i-1} + |b''_i| b'_i\,.$$

Quindi, essendo un insieme di n generatori di uno spazio vettoriale di dimensione n, per il Teorema 5.4.13, essi sono una base di V.

Passiamo alla ortonormalità.

1) $\langle b'_i, b'_i\rangle = \langle \frac{b''_i}{|b''_i|}, \frac{b''_i}{|b''_i|}\rangle = \frac{\langle b''_i, b''_i\rangle}{|b''_i|^2} = \frac{\langle b''_i, b''_i\rangle}{\langle b''_i, b''_i\rangle} = 1$.

2) $\langle b'_i, b'_j\rangle = 0$ se $i \neq j$.

Infatti:

$$\langle b'_2, b'_1\rangle = \frac{1}{|b''_2|}(\langle b_2, b'_1\rangle - \langle b_2, b'_1\rangle\langle b'_1, b'_1\rangle) = 0\,,$$

perché $\langle b'_1, b'_1\rangle = 1$.

Supponiamo poi che, fissato h, con $1 < h < n$, $\langle b'_h, b'_k\rangle = 0$ per ogni $k < h$. Dimostriamo che $\langle b'_{h+1}, b'_k\rangle = 0$ per ogni $k < h + 1$. Così facendo, la

dimostrazione è completata.

$$\langle b''_{h+1}, b'_k \rangle = \langle b_{h+1}, b'_k \rangle - \langle b_{h+1}, b'_1 \rangle \langle b'_1, b'_k \rangle - \cdots - \langle b_{h+1}, b'_h \rangle \langle b'_h, b'_k \rangle .$$

Per ipotesi induttiva, tutti i $\langle b'_j, b'_k \rangle$ sono nulli se $j \neq k$ perché sia k che j sono minori o uguali a h. Della espressione precedente rimangono solo il primo termine e quello relativo a b'_k. Si ha:

$$\langle b''_{h+1}, b'_k \rangle = \langle b_{h+1}, b'_k \rangle - \langle b_{h+1}, b'_k \rangle \langle b'_k, b'_k \rangle = 0 ,$$

poiché $\langle b'_k, b'_k \rangle = 1$. Ne segue che $\langle b'_{h+1}, b'_k \rangle = \frac{1}{|b''_{h+1}|} \langle b''_{h+1}, b'_k \rangle = 0$, per ogni $k < h + 1$.

Notiamo che l'ipotesi che $\langle \, , \rangle$ sia definito positivo viene utilizzata per poter dividere il vettore non nullo b''_i per la sua norma che, se $\langle \, , \rangle$ è definito positivo, è sicuramente diversa da zero. Per inciso osserviamo che se b''_i fosse 0, allora b_i potrebbe essere ottenuto come una combinazione lineare dei b'_j con $j < i$, che a loro volta sono ottenuti come combinazioni dei b_j, con $j < i$, ma ciò non è possibile perché per ipotesi $b_1, \dots, b_n$ sono linearmente indipendenti.

8.3.2 Segnatura di un prodotto scalare

Sia $\langle \, , \rangle$ un prodotto scalare (qualsiasi) su V.

Proposizione 8.3.10 *Sia $\mathcal{B}$ una base ortogonale di $\langle \, , \rangle$. Se $n^0(\mathcal{B})$ è il numero di vettori della base $\mathcal{B}$ tali che $\langle b_i, b_i \rangle = 0$, allora $n^0(\mathcal{B}) = \dim V_0$, il nucleo di $\langle \, , \rangle$.*

Dim. A meno di riordinare la base $\mathcal{B}$, possiamo supporre:

$$\begin{aligned} \langle b_i, b_i \rangle = 0 \qquad & \text{per } 1 \leq i \leq n^0(\mathcal{B}), \\ \langle b_i, b_i \rangle \neq 0 \qquad & \text{per } n^0(\mathcal{B}) < i \leq n . \end{aligned}$$

Poiché $\mathcal{B}$ era ortogonale, per $1 \leq i \leq n^0(\mathcal{B})$, $\langle b_i, b_k \rangle = 0$, $1 \leq k \leq n$.

Allora $b_i \in V_0$.

Essi sono anche dei generatori di V_0. Infatti sia $v \in V_0$, $v = \lambda_1 b_1 + \cdots + \lambda_n b_n$. Per $k > n^0(\mathcal{B})$ si ha: $0 = \langle v, b_k \rangle = \lambda_k \langle b_k, b_k \rangle$, ma poiché $\langle b_k, b_k \rangle \neq 0$, allora $\lambda_k = 0$. Dunque v è generato da $b_1, \dots, b_{n^0(\mathcal{B})}$. □

Quindi $n^0(\mathcal{B})$ non dipende dalla scelta della base $\mathcal{B}$, ma solo da $\langle\, ,\rangle$. Il numero $n^0 = n^0(\mathcal{B})$ è detto *indice di nullità* del prodotto scalare $\langle\, ,\rangle$.

Sia $\mathcal{B}$ una base ortogonale di $\langle\, ,\rangle$. Analogamente possiamo definire:
$n^+(\mathcal{B})$ è il numero di vettori della base $\mathcal{B}$ tali che $\langle b_i, b_i\rangle > 0$,
$n^-(\mathcal{B})$ è il numero di vettori della base $\mathcal{B}$ tali che $\langle b_i, b_i\rangle < 0$.

Vale il seguente teorema:

Teorema 8.3.11 (Sylvester) *Se $\mathcal{B}$ e $\mathcal{B}'$ sono due diverse basi ortogonali per il prodotto scalare $\langle\, ,\rangle$ di V, allora:*

$$n^+(\mathcal{B}) = n^+(\mathcal{B}')\,.$$

Quindi il numero $n^+(\mathcal{B})$ non dipende da $\mathcal{B}$. Viene quindi indicato semplicemente con n^+ e detto *indice di positività* di $\langle\, ,\rangle$.

Inoltre, poiché $n^+(\mathcal{B}) + n^-(\mathcal{B}) + n^0(\mathcal{B}) = n$, la dimensione di V, anche $n^-(\mathcal{B})$ non dipende da $\mathcal{B}$. Il numero $n^- = n^-(\mathcal{B})$ è detto *indice di negatività* di $\langle\, ,\rangle$.

La coppia (n^+, n^-) si dice anche *segnatura* del prodotto scalare.

8.4 Endomorfismi unitari

Sia $\langle\, ,\rangle$ un prodotto scalare definito positivo. Come abbiamo detto in precedenza, il numero reale positivo $|v| = \sqrt{\langle v, v\rangle}$ è detto norma di v o anche lunghezza di v.

Si può dimostrare che:
1) $|\lambda v| = |\lambda||v|$ per ogni $v \in V$, $\lambda \in \mathbb{R}$.
2) $|v| \geq 0$ e $|v| = 0$ se e solo se $v = 0$.
3) $|v + w| \leq |v| + |w|$.

Dimostriamo la 3). Premettiamo il seguente

Lemma 8.4.1 (Disuguaglianza di Schwarz) *Siano $v, w \in V$. Allora:*

$$|\langle v, w\rangle| \leq |v||w|\,.$$

Dim. La disuguaglianza da dimostrare è equivalente alla:

$$\langle v, w\rangle^2 \leq \langle v, v\rangle\langle w, w\rangle\,.$$

Se $w = 0$, vale banalmente l'uguaglianza.

Se $w \neq 0$, allora $\langle w, w\rangle > 0$. In questo caso poniamo:

$$a = \langle w, w\rangle\,, \quad b = -\langle v, w\rangle\,.$$

Abbiamo:

$$\begin{aligned} 0 &\leq \langle av + bw, av + bw\rangle = a^2\langle v, v\rangle + 2ab\langle v, w\rangle + b^2\langle w, w\rangle = \\ &= \langle w, w\rangle^2\langle v, v\rangle - 2\langle w, w\rangle\langle v, w\rangle^2 + \langle v, w\rangle^2\langle w, w\rangle\,. \end{aligned}$$

Dividendo per $\langle w, w\rangle$ si ha:

$$0 \leq \langle w, w\rangle\langle v, v\rangle - \langle v, w\rangle^2\,.$$

□

Adesso abbiamo:

$$\begin{aligned} |v + w|^2 &= \langle v + w, v + w\rangle = |v|^2 + |w|^2 + 2\langle v, w\rangle \\ &\leq |v|^2 + |w|^2 + 2|\langle v, w\rangle| \leq |v|^2 + |w|^2 + 2|v||w| \\ &= (|v| + |w|)^2\,. \end{aligned}$$

Da cui:

$$|v + w| \leq |v| + |w|\,.$$

D'ora in avanti, salvo esplicito avvertimento, tutti gli spazi vettoriali saranno muniti di un prodotto scalare $\langle\,,\rangle$ definito positivo.

Proposizione 8.4.2 *La funzione $d : V \times V \to \mathbb{R}$, $d(x, y) = |x - y|$ è una metrica su V, cioè una funzione che verifica le seguenti tre proprietà:*
1) $d(x, y) \geq 0$ per ogni $x, y \in V$ e $d(x, y) = 0$ se e solo se $x = y$.
2) $d(x, y) = d(y, x)$ per ogni $x, y \in V$.
3) $d(x, z) \leq d(x, y) + d(y, z)$ per ogni $x, y, z \in V$.

Se $V = \mathbb{R}^2$ (oppure $\mathbb{R}^3$) e $\langle\,,\rangle$ è il prodotto scalare canonico, allora d coincide con la distanza ordinaria.

Definizione 8.4.3 *Un endomorfismo $f : V \to V$ si dice* unitario *se:*

$$\langle f(x), f(y)\rangle = \langle x, y\rangle \quad \forall x, y \in V\,.$$

Osservazione 8.4.4 *Un endomorfismo unitario è una isometria, cioè una applicazione che lascia inalterata la distanza:*

$$d(x, y) = d(f(x), f(y)) .$$

Vale anche un "quasi inverso" della Osservazione 8.4.4.

Teorema 8.4.5 *Se* $\sigma : V \to V$ *è una isometria per la distanza definita su* V *da un prodotto scalare definito positivo* $\langle\, , \rangle$ *ed inoltre* $\sigma(0) = 0$*. Allora* σ *è un endomorfismo unitario.*

La dimostrazione di questo fatto è alquanto laboriosa ma poco istruttiva e pertanto viene omessa.

La seguente proposizione è utilizzata quando occorre verificare che un certo endomorfismo è unitario.

Proposizione 8.4.6 *Sia* f *un endomorfismo di* V*. Le seguenti proprietà sono equivalenti:*

1) f *è unitario,*

2) $\langle f(x), f(x)\rangle = \langle x, x\rangle$ *per ogni* $x \in V$*,*

3) $\langle f(x), f(x)\rangle = \langle x, x\rangle$ *per ogni* $x \in V$ *tale che* $\langle x, x\rangle = 1$*.*

Dim. Chiaramente 1) implica 2) e 2) implica 3).

Dimostriamo che 3) implica 2).

Sia $x \in V$. Se $x = 0$, chiaramente la 2) vale.

Se $x \neq 0$, per ipotesi, in quanto $\langle \frac{x}{|x|}, \frac{x}{|x|}\rangle = 1$, si ha:

$$\langle \frac{x}{|x|}, \frac{x}{|x|}\rangle = \langle f\Big(\frac{x}{|x|}\Big), f\Big(\frac{x}{|x|}\Big)\rangle .$$

Poiché f è lineare, $f\big(\frac{x}{|x|}\big) = \frac{f(x)}{|x|}$, e quindi la tesi:

$$\langle f(x), f(x)\rangle = \langle x, x\rangle \quad \forall x \in V .$$

Dimostriamo che 2) implica 1).

Per ipotesi $\langle x + y, x + y\rangle = \langle f(x + y), f(x + y)\rangle$ per ogni $x, y \in V$.

Sviluppando, si ha:

$$\begin{aligned} &\langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle = \\ &\quad = \langle f(x), f(x)\rangle + \langle f(x), f(y)\rangle + \langle f(y), f(x)\rangle + \langle f(y), f(y)\rangle . \end{aligned}$$

Utilizzando ancora l'ipotesi che $\langle x, x\rangle = \langle f(x), f(x)\rangle$, per ogni x, otteniamo $2\langle x, y\rangle = 2\langle f(x), f(y)\rangle$, da cui la tesi:

$$\langle x, y\rangle = \langle f(x), f(y)\rangle \quad \forall x, y \in V\,,$$

cioè f è unitario. □

Gli endomorfismi unitari sono caratterizzati dalla seguente proprietà:

Proposizione 8.4.7 *Supponiamo di fissare su V una base ortonormale $\mathcal{B}$ per il prodotto scalare definito positivo $\langle\,,\rangle$. Sia f un endomorfismo.*

Se A è la matrice associata a f nella base $\mathcal{B}$, si ha:

f è unitario se e solo se A verifica la relazione

$${}^tAA = I\,.$$

Dim. Nella base $\mathcal{B}$ la matrice che esprime $\langle\,,\rangle$ è la matrice identità I.

Se $(x_1, \ldots, x_n)$ e $(y_1, \ldots, y_n)$ sono le coordinate di v e w, rispettivamente,

$$\langle v, w\rangle = (x_1, \ldots, x_n)\cdot\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\,,$$

$$f(v) = A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\,, \quad f(w) = A\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\,,$$

$$\langle f(v), f(w))\rangle = (x_1, \ldots, x_n)\,{}^tA\cdot A\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\,.$$

Quindi, se ${}^tAA = I$, si ha $\langle v, w\rangle = \langle f(v), f(w)\rangle$ e f è unitario.

Viceversa, supponiamo f unitario.

Se c_{ij} è l'elemento di posto (i, j) della matrice ${}^tAA$, si ha:

$$c_{ij} = e_i\,{}^tA\cdot Ae_j = \langle e_i, e_j\rangle = \delta_{ij}\,.$$

Quindi ${}^tA\cdot A = I$. □

Definizione 8.4.8 *Una matrice $A \in \mathbf{M}_n(\mathbb{R})$ si dice* ortogonale *se:*

$${}^tA\cdot A = A\cdot{}^tA = I\,.$$

Osserviamo che, se A è ortogonale, A è invertibile e ${}^tA = A^{-1}$.

Proposizione 8.4.9 *L'insieme delle matrici ortogonali di ordine n è un gruppo con l'operazione di prodotto riga per colonna. Tale gruppo si dice* gruppo ortogonale *e si indica con $O(n)$.*

Dim.
1) $I_n \in O(n)$, ovviamente.
2) Se $A, B \in O(n)$, anche $AB \in O(n)$.
Infatti, ${}^t(AB)(AB) = {}^tB({}^tA \cdot A)B = {}^tB \cdot B = I$.
3) Se $A \in O(n)$, allora A è invertibile e $A^{-1} \in O(n)$.
Infatti, abbiamo già osservato che ${}^tA = A^{-1}$. Allora:

$$ {}^t(A^{-1})A^{-1} = {}^t({}^tA)\,{}^tA = A\,{}^tA = I\,. $$

□

Osservazione 8.4.10 *Se $A \in O(n)$, allora* $\det A = 1$ *oppure* $\det A = -1$.
Infatti $\det({}^tA \cdot A) = \det({}^tA)\det A = \det A \det A = (\det A)^2$. *Ma anche* $\det({}^tA \cdot A) = \det I = 1$. *Pertanto* $(\det A)^2 = 1$. *Da cui la tesi.*

Osservazione 8.4.11 *Se M è la matrice di passaggio delle coordinate da una base ortonormale $\mathcal{B}$ ad un'altra base ortonormale $\mathcal{B}'$, allora $M \in O(n)$.*
Infatti sia in base $\mathcal{B}$ che in base $\mathcal{B}'$, la matrice associata a $\langle\,,\rangle$ è I. Per la formula del cambiamento di base (8.2), si ha che:

$$ I = {}^tM \cdot I \cdot M = {}^tM \cdot M\,. $$

Vogliamo studiare ora come è fatto $O(2)$.
Sia $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ una matrice di ordine 2. $A \in O(2)$ se e solo se:

$$ \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. $$

Quindi a, b, c, d debbono verificare il seguente sistema, ottenuto uguagliando termine a termine le due matrici.

$$ \begin{cases} a^2 + c^2 = 1\,, \\ ab + dc = 0\,, \\ b^2 + d^2 = 1\,. \end{cases} $$

Supponiamo che $\det A = 1$.

Alle tre equazioni del sistema se ne aggiunge una quarta: $ad - bc = 1$.
Per la prima equazione del sistema, esiste un unico $\theta \in \mathbb{R}$ tale che: $0 \leq \theta < 2\pi$ e inoltre $a = \cos\theta$ e $c = \sin\theta$.
Per la terza equazione del sistema, esiste un unico $\omega \in \mathbb{R}$ tale che: $0 \leq \omega < 2\pi$ e inoltre $b = \sin\omega$ e $d = \cos\omega$.
La seconda equazione del sistema diventa: $\cos\theta\sin\omega + \sin\theta\cos\omega = 0$, ovvero $\sin(\theta + \omega) = 0$.
La quarta equazione, quella che esprime $\det A = 1$, diventa: $\cos\theta\cos\omega - \sin\theta\sin\omega = 1$, ovvero $\cos(\theta + \omega) = 1$. Questo implica che $\theta + \omega = 0$ oppure $\theta + \omega = 2\pi$. In ogni caso, sia che $\omega = -\theta$ che $\omega = 2\pi - \theta$, poiché $\sin(-\theta) = \sin(2\pi - \theta) = -\sin\theta$ e $\cos(-\theta) = \cos(2\pi - \theta) = \cos\theta$, si ha che:

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad \text{con } 0 \leq \theta < 2\pi\,.$$

Se $\det A = -1$, restando invariate le considerazioni fatte più sopra per le prime tre equazioni, la quarta equazione diventa: $\cos\theta\cos - \sin\theta\sin\omega = -1$ ovvero $\cos(\theta + \omega) = -1$. Questo implica che $\theta + \omega = \pi$ oppure $\theta + \omega = 3\pi$. In ogni caso, sia che $\omega = \pi - \theta$ che $\omega = 2\pi + (\pi - \theta)$, poiché $\sin(\pi - \theta) = \sin\theta$ e $\cos(\pi - \theta) = -\cos\theta$ si ha che:

$$A = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}, \qquad \text{con } 0 \leq \theta < 2\pi\,.$$

Osservazione 8.4.12 *Sia f un endomorfismo unitario. Seλ è un autovalore di f, allora $\lambda = 1$ oppure $\lambda = -1$.*

Infatti, se x è un autovettore di f relativo a λ,

$$\langle x, x\rangle = \langle f(x), f(x)\rangle = \langle \lambda x, \lambda x\rangle = \lambda^2\langle x, x\rangle\,,$$

da cui, essendo $\langle x, x\rangle \neq 0$, si ricava $\lambda^2 = 1$, cioè $\lambda = \pm 1$.

Utilizziamo questa osservazione per descrivere $O(3)$.

Sia $A \in O(3)$. Poiché A ha ordine dispari essa ammette un autovalore reale e siccome A rappresenta, in una qualche base $\mathcal{B}$ ortonormale di V, un endomorfismo unitario f, tale autovalore $\lambda = \pm 1$. Supponiamo che x sia un autovettore relativo a λ con norma $|x| = 1$.

Possiamo scegliere x come primo vettore di un'altra base ortonormale $\mathcal{B}'$. In questa nuova base f è rappresentato dalla matrice A' tale che: $A' = M^{-1}AM$. Siccome però M è la matrice di passaggio da una base ortonormale ad un'altra base ortonormale $M \in O(3)$ e quindi: $A' = {}^tMAM$. Allora anche $A' \in O(3)$. Siccome, inoltre il primo elemento della nuova base è un autovettore di A, A' è della forma:

$$\begin{pmatrix} \lambda & \mu & \nu \\ 0 & a & b \\ 0 & c & d \end{pmatrix},$$

dove $\lambda = \pm 1$, $A' \in O(3)$. Ciò implica che ${}^tA'A' = I$:

$$\begin{pmatrix} \lambda & 0 & 0 \\ \mu & a & b \\ \nu & b & d \end{pmatrix} \begin{pmatrix} \lambda & \mu & \nu \\ 0 & a & b \\ 0 & c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Eseguendo i calcoli si ha: $\lambda\mu = 0$ e $\lambda\nu = 0$, quindi $\mu = \nu = 0$. Inoltre:

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \text{ cioè } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in O(2).$$

In definitiva, se $A \in O(3)$, allora esiste $M \in O(3)$ tale che $A' = {}^tMAM$ dove

$$A' = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \text{ oppure } A' = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & \sin\theta & -\cos\theta \end{pmatrix}$$

con $\lambda = \pm 1$ e $0 \leq \theta < 2\pi$.

8.5 Endomorfismi simmetrici

Sia $\langle\, , \rangle$ un prodotto scalare non degenere su V e sia $\phi : V \to V$ un endomorfismo di V.

Definizione 8.5.1 *Un endomorfismo $\psi : V \to V$ si dice* aggiunto *di ϕ se:*

$$\langle \phi(x), y \rangle = \langle x, \psi(y) \rangle \quad \forall x, y \in V.$$

Un endomorfismo ϕ si dice autoaggiunto *o* simmetrico *se coincide con il proprio aggiunto, ossia:*

$$\langle \phi(x), y \rangle = \langle x, \phi(y) \rangle \quad \forall x, y \in V.$$

Osservazione 8.5.2 *Dato ϕ, il suo aggiunto esiste sempre. Infatti fissiamo su V una base ortogonale $\mathcal{B}$. In tale base sia A la matrice associata a ϕ, e sia M la matrice associata al prodotto scalare $\langle\,,\rangle$.*

Se in tale base $(x_1,\ldots,x_n)$ e $(y_1,\ldots,y_n)$ sono le coordinate di v e w rispettivamente, allora

$$\langle\phi(v),w\rangle = (x_1,\ldots,x_n)\,{}^tAM\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}.$$

Se ψ è un endomorfismo di V tale che $\langle\phi(v),w\rangle = \langle v,\psi(w)\rangle$, la matrice A' che esprime ψ nella base $\mathcal{B}$ dovrà verificare l'uguaglianza:

$$(x_1,\ldots,x_n)\,{}^tAM = (x_1,\ldots,x_n)MA'\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}\qquad \forall x,y\in V\,.$$

Ciò accade se e solo se ${}^tAM = MA'$. Poiché $\langle\,,\rangle$ è non degenere, sappiamo, per la Proposizione 8.2.5, che $\det M \neq 0$.

Quindi l'aggiunto di ϕ è l'endomorfismo ψ che nella base ortogonale in cui ϕ è espresso dalla matrice A, è espresso dalla matrice A' con: $A' = M^{-1}\,{}^tAM$, dove M è la matrice che rappresenta $\langle\,,\rangle$.

Osservazione 8.5.3 *Se $\langle\,,\rangle$ è definito positivo, allora in una base ortonormale la matrice che esprime $\langle\,,\rangle$ è l'identità I. Allora in tale base l'aggiunto di ϕ è espresso dalla matrice $A' = {}^tA$.*

In particolare ϕ è autoaggiunto se e solo se è espresso, in una base ortonormale, da una matrice simmetrica.

Vogliamo a questo punto dimostrare il seguente importante risultato.

Teorema 8.5.4 *Ogni matrice simmetrica ha tutti gli autovalori reali e inoltre è diagonalizzabile.*

Per poterlo dimostrare abbiamo bisogno di alcune definizioni sugli spazi complessi.

Definizione 8.5.5 *Sia V un $\mathbb{C}$-spazio vettoriale. Una applicazione $h : V\times V\to\mathbb{C}$ è un* prodotto hermitiano *se valgono le seguenti proprietà:*

1. $h(\lambda z_1 + \mu z_2, w) = \lambda h(z_1, w) + \mu h(z_2, w)$,

2. $h(z_1, z_2) = \overline{h(z_2, z_1)}$ *per ogni* $z_1, z_2, w \in V$, $\lambda, \mu \in \mathbb{C}$.

Quindi $h(z, \lambda w_1 + \mu w_2) = \overline{\lambda} h(z, w_1) + \overline{\mu} h(z, w_2)$.

Osservazione 8.5.6 *Se h è un prodotto hermitiano, allora $h(z, z) \in \mathbb{R}$.*

Definizione 8.5.7 *Un prodotto hermitiano h si dice definito positivo se: $h(z, z) \geq 0$ e $h(z, z) = 0$ se e solo se $z = 0$.*

Se si fissa nel $\mathbb{C}$-spazio V una base, in perfetta analogia con quanto visto all'inizio della sezione 8.2, si può associare al prodotto hermitiano h una matrice H. Solo che, mentre le matrici associate ai prodotti scalari reali erano simmetriche, in questo caso, la matrice H avrà la seguente proprietà:

$$^tH = \overline{H}\,. \tag{8.3}$$

Definizione 8.5.8 *Una matrice H che coincide con la propria trasposta coniugata $^t\overline{H}$, si dice* hermitiana.

Analogamente al caso reale, anche per i prodotti hermitiani è possibile dare le definizioni di basi ortogonali o ortonormali e di endomorfismi simmetrici.

Osservazione 8.5.9 *La matrice associata ad un endomorfismo simmetrico in una base ortonormale, è una matrice hermitiana.*

Proposizione 8.5.10 *Sia A una matrice simmetrica (reale). Allora tutti i suoi autovalori sono reali.*

Dim. Poniamo su $\mathbb{C}^n$ il prodotto hermitiano canonico h, quello che nella base canonica è espresso dalla matrice I. Siccome A è simmetrica e reale, essa è hermitiana e dunque definisce nella base canonica di $\mathbb{C}^n$ un endomorfismo simmetrico ϕ. Sia λ un autovalore di ϕ (quindi di A) e x un autovettore di ϕ relativo a λ. Si ha: $h(\phi(x), x) = h(x, \phi(x))$, perché ϕ è simmetrico. Poiché $\phi(x) = \lambda x$, si ha: $\lambda h(x, x) = h(\lambda x, x) = h(x, \lambda x) = \overline{\lambda} h(x, x)$. Inoltre $h(x, x) > 0$, essendo $x \neq 0$. Quindi $\lambda = \overline{\lambda}$, cioè $\lambda \in \mathbb{R}$. $\square$

Vogliamo utilizzare questo risultato per dimostrare la possibilità di diagonalizzare simultaneamente due prodotti scalari, uno dei quali almeno definito positivo.

Teorema 8.5.11 (Teorema spettrale) *Sia $(\,,\,)$ un prodotto scalare dello spazio vettoriale reale V e sia $\langle\,,\,\rangle$ un prodotto scalare di V definito positivo. Esiste una base di V ortonormale per $\langle\,,\,\rangle$ e ortogonale per $(\,,\,)$.*

Dim. Dimostriamolo per induzione su n, la dimensione di V.

Se $n = 1$, l'enunciato è vero. Ogni vettore non nullo v di V tale che $\langle v, v\rangle = 1$ è una base con i requisiti richiesti.

Supponiamo che l'enunciato valga per $n-1$ e dimostriamo che vale anche per n.

Se $(\,,\,)$ è identicamente nullo, ogni base di V è ortogonale per $(\,,\,)$ e quindi basta scegliere una base ortonormale di $\langle\,,\,\rangle$ per aver finito.

Supponiamo allora che $(\,,\,)$ non sia identicamente nullo e sia A la matrice associata a $(\,,\,)$ in una base $\mathcal{B}$ di V ortonormale per $\langle\,,\,\rangle$. In queste coordinate, la matrice associata a $\langle\,,\,\rangle$ è I, mentre A è semplicemente simmetrica e diversa da 0.

Se $v, w \in V$ con coordinate, in base $\mathcal{B}$, rispettivamente $(x_1, \dots, x_n)$ e $(y_1, \dots, y_n)$, si ha

$$(v, w) = (x_1, \dots, x_n)A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \langle v, \phi(w)\rangle\,,$$

dove ϕ è l'endomorfismo simmetrico di V espresso, in base $\mathcal{B}$, dalla matrice simmetrica A. Sia λ un autovalore reale di ϕ, che è anche autovalore di A (questo esiste per la Proposizione 8.5.10). Sia u un autovettore di ϕ relativo a λ. Allora possiamo scrivere:

$$(v, u) = \langle v, \phi(u)\rangle = \langle v, \lambda u\rangle = \lambda\langle v, u\rangle \quad \forall v \in V\,. \tag{8.4}$$

Se poniamo:

$$\{u\}^{\perp} = \{v \in V \mid (v, u) = 0\}\,, \quad \{u\}^{perp} = \{v \in V \mid \langle v, u\rangle = 0\}$$

abbiamo che, per la relazione (8.4), $\{u\}^{perp} \subseteq \{u\}^{\perp}$.

Poiché $\langle\,,\,\rangle$ è definito positivo, $\{u\}^{perp}$ ha dimensione $n-1$. Per ipotesi induttiva, sia $\mathcal{B}'$ una base di $\{u\}^{perp}$ che sia ortonormale per $\langle\,,\,\rangle$ e ortogonale per $(\,,\,)$. Se poniamo $u_1 = \frac{u}{|u|}$, allora $\mathcal{C} = \{u_1\} \cup \mathcal{B}'$ è la base cercata. □

Come conseguenza, dati due prodotti scalari di cui uno definito positivo, esiste sempre una base in cui le matrici associate ai due prodotti scalari siano entrambe diagonali.

Corollario 8.5.12 *Una matrice simmetrica è sempre diagonalizzabile.*

Dim. Consideriamo $\mathbb{R}^n$ con la base canonica e munito del prodotto scalare canonico $\langle\, , \rangle$, che è definito positivo. Sia A la matrice simmetrica data e sia $(\, ,)$ il prodotto scalare definito in $\mathbb{R}^n$ da A nella base canonica. Per il Teorema 8.5.11 sappiamo che esiste una base $\mathcal{B}$ ortonormale per $\langle\, , \rangle$ e ortogonale per $(\, ,)$. Sia M la matrice di passaggio dalla base $\mathcal{B}$ nella base canonica. Nella base $\mathcal{B}$, $(\, ,)$ è espresso da una matrice D diagonale, legata alla matrice A dalla relazione $D = {}^tMAM$, per la formula del cambiamento di base (8.2). Poiché però M è una matrice di passaggio da una base ortonormale ad un'altra base ortonormale per $\langle\, , \rangle$, per l'Osservazione 8.4.11, $M \in O(n)$ e dunque ${}^tM = M^{-1}$. Allora $D = M^{-1}AM$ e quindi A e D sono simili. □

8.6 Un criterio per i prodotti scalari definiti positivi

Osservazione 8.6.1 *Sia $\mathcal{B}$ una base ortogonale per il prodotto scalare $\langle\, , \rangle$. In questa base, la matrice associata a $\langle\, , \rangle$ è una matrice diagonale, poniamo $D = diag(d_1, \ldots, d_n)$. Allora $\langle\, , \rangle$ è definito positivo se e solo se $d_i > 0$ per ogni i.*

Infatti se $(x_1, \ldots, x_n)$ sono le coordinate del generico vettore v di V in base $\mathcal{B}$, si ha che:

$$\langle v, v\rangle = (x_1, \ldots, x_n)D \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = d_1x_1^2 + \cdots + d_nx_n^2 \,.$$

Se $d_i > 0$ per ogni i, allora è chiaro che $\langle v, v\rangle \geq 0$ e $\langle v, v\rangle = 0$ se e solo se $x_i = 0$ per ogni i, cioè $v = 0$.

Viceversa, se anche un solo d_i (ad esempio d_1) fosse minore o uguale a zero,si avrebbe $d_1 = \langle b_1, b_1\rangle \leq 0$, dove b_1 è il primo elemento di $\mathcal{B}$. Quindi $\langle\, , \rangle$ non sarebbe definito positivo.

Osservazione 8.6.2 *Se A è la matrice che esprime $\langle\, ,\rangle$ in una certa base, sappiamo che esiste una matrice ortogonale M tale che $D = {}^tMAM$ è diagonale e A e D sono simili. Se $D = diag(d_1, \dots, d_n)$, la matrice D, e quindi A, definiscono un prodotto scalare definito positivo se e solo se tutti i d_i sono positivi. Ma i d_i sono gli autovalori di D e quindi di A. Quindi, affinché A esprima un prodotto scalare definito positivo, deve avere tutti gli autovalori positivi (già sappiamo che, essendo A simmetrica, i suoi autovalori sono tutti reali).*

Per verificare che una matrice simmetrica A abbia tutti gli autovalori positivi, basta applicare la regola di Cartesio al polinomio caratteristico $p_A(t)$ di A, cioè occorre che la sequenza dei coefficienti di $p_A(t)$ abbia esattamente n variazioni di segno.

Questo criterio è equivalente al seguente: *la matrice simmetrica A definisce un prodotto scalare definito positivo se i determinanti di tutti i minori principali di A sono positivi.*

(Un minore è detto principale se la sua diagonale principale è un sottoinsieme delle diagonale principale di A).

L'equivalenza di questi criteri sta nel fatto che il coefficiente del termine di grado k di $p_A(t)$ è dato dalla somma di tutti i determinanti dei minori principali di A di ordine $n - k$, moltiplicata per $(-1)^{n-k}$.

Più precisamente vale il seguente:

Criterio 8.6.3 (criterio di Sylvester) *Una matrice simmetrica A di ordine n definisce un prodotto scalare definito positivo se e solo se*

$$\Delta_r(A) > 0 \, per\ ogni\ r,\ con\ 1 \leq r \leq n, \tag{8.5}$$

dove con $\Delta_r(A)$ indichiamo il determinante del minore di A ottenuto prendendo le prime r righe e le prime r colonne di A.

8.7 Il teorema di scomposizione polare

Premettiamo il seguente Lemma al risultato principale.

Lemma 8.7.1 *Sia A matrice simmetrica (reale) definita positiva, cioè con tutti gli autovalori positivi. Allora esiste (unica) B simmetrica definita positiva tale che $A = B^2$.*

Dim. La matrice A è diagonalizzabile per il teorema spettrale. Allora $N^{-1}AN = D$ con N matrice ortogonale e D diagonale. Sulla diagonale di D compaiono gli autovalori di A e, essendo questa definita positiva, essi sono tutti positivi. Dunque esiste R diagonale definita positiva tale che $R^2 = D$.

Abbiamo in conclusione:

$$A = NDN^{-1} = NR^2N^{-1} = NRIRN^{-1} = (NRN^{-1})(NRN^{-1})\,.$$

La matrice NRN^{-1} è simmetrica definita positiva.

La dimostrazione dell'unicità è lasciata per esercizio. □

Teorema 8.7.2 (scomposizione polare) *Ogni matrice invertibile si può scrivere (in modo unico) come prodotto di una matrice ortogonale e di una matrice simmetrica definita positiva.*

Dim. Sia A una matrice invertibile. Si mostra subito che la matrice ${}^tAA$ è simmetrica e definita positiva. Allora, per il Lemma 8.7.1, esiste S, simmetrica definita positiva, tale che ${}^tAA = S^2$, cioè $A = ({}^tA)^{-1}S^2$. Basta allora osservare che la matrice $({}^tA)^{-1}S$ è ortogonale.

La dimostrazione dell'unicità è lasciata per esercizio. □

8.8 Esercizi

Diagonalizzazione di matrici simmetriche

1) Determinare per ciascuna delle seguenti matrici (reali simmetriche) una matrice ortogonale che la diagonalizzi e si scriva la matrice diagonale corrispondente:

$$A = \begin{pmatrix} 5 & 2 & 3 \\ 2 & -1 & 0 \\ 3 & 0 & 1 \end{pmatrix}, \tag{8.6}$$

$$B = \begin{pmatrix} 1/2 & 3/2 & 0 \\ 3/2 & 1/2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \tag{8.7}$$

$$C = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \tag{8.8}$$

$$D = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 0 & 0 \\ -3 & 0 & 9 \end{pmatrix}, \tag{8.9}$$

$$E = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}. \tag{8.10}$$

Ortogonalizzazione di Gram-Schmidt e matrici ortogonali

2) Usando il processo di ortogonalizzazione di Gram-Schmidt, provare che ogni matrice reale invertibile si può scrivere come prodotto di una matrice ortogonale e di una triangolare superiore (scomposizione polare).

3) Data una matrice ortogonale A di ordine n, provare che, se $\det A = (-1)^{n+1}$, allora 1 è autovalore di A.

Capitolo 9

ELEMENTI DI GEOMETRIA ANALITICA

9.1 Coordinate cartesiane sulla retta e nel piano

Su una retta fissiamo un punto O ed un punto A distinto da O. Se scegliamo il segmento OA come unità di misura, possiamo far corrispondere ad ogni punto P della retta un numero reale nel seguente modo.

Se P appartiene alla stessa semiretta di origine O che contiene A, allora a P si fa corrispondere il numero reale positivo dato dalla misura del segmento OP rispetto all'unità di misura OA.

Se P appartiene alla semiretta opposta a quella che contiene A, allora a P si fa corrispondere il numero reale opposto della misura di OP, sempre rispetto ad OA.

Si può dimostrare che tale corrispondenza tra punti della retta e $\mathbb{R}$ è biunivoca (anzi gli assiomi di Euclide che definiscono la retta sono stati via via integrati, soprattutto nel secolo scorso, da Cantor e Dedekind, ad esempio, in modo tale da ottenere questa corrispondenza biunivoca, quindi ogni speculazione metafisica a tale proposito va decisamente ridimensionata).

Al punto O corrisponde 0, al punto A corrisponde 1; ai punti sulla semiretta OA i numeri reali positivi e all'altra semiretta i numeri negativi.

In questo contesto, se al punto M corrisponde in numero x e al punto N il numero y, la lunghezza del segmento MN sarà $|x-y|$.

Vogliamo porre, ora, delle coordinate anche nel piano euclideo.

Scegliamo allora, nel piano, due rette perpendicolari tra loro, r_1 e r_2, e sia O il punto di intersezione.

Scegliamo su ciascuna retta un punto, rispettivamente A' e A'', tali che i segmenti OA' e OA'' siano di ugual lunghezza.

Procedendo come fatto prima, su ciascuna retta risulta fissata una corrispondenza biunivoca con $\mathbb{R}$.

Sia ora P un punto appartenente al piano (eventualmente anche ad una delle due rette). Da P conduciamo le perpendicolari a r_1 e r_2 e siano P_1 e P_2, rispettivamente, i punti di intersezione di tali perpendicolari con r_1 e r_2. Poiché P_1 e P_2 appartengono rispettivamente a r_1 e r_2, a ciascuno di essi corrisponde un numero reale. Siano x_1 e x_2 tali numeri reali.

Con questa costruzione, una volta fissate le due rette, l'unità di misura per i segmenti e l'origine, ad ogni punto P del piano si fa corrispondere una coppia ordinata di numeri reali (x_1, x_2).

La coppia di rette con la scelta dell'origine e delle unità di misura, viene detto *sistema di riferimento cartesiano ortogonale monometrico* (con il chiaro riferimento di ciascun termine: cartesiano perché si scelgono come riferimento due rette; ortogonale perché le rette vengono scelte perpendicolari; monometrico perché l'unità di misura è la stessa su entrambe le rette. Ovviamente tali scelte sono arbitrarie e in alcuni contesti possono essere modificate, scegliendo sistemi curvilinei o rette oblique o unità differenti su ciascuna retta).

La corrispondenza $P \to (x_1, x_2)$ tra il piano e $\mathbb{R}^2$ è biunivoca e quindi ogni punto è ben individuato dalla coppia delle sue coordinate.

Il numero x_1 è detto *ascissa* di P e x_2 è detto *ordinata*.

Per associare alla coppia (x_1, x_2) il corrispondente punto del piano si procede così.

Sia P_1 il punto sulla retta r_1 (detta asse delle ascisse) che corrisponde al numero reale x_1 e si tracci da P_1 la perpendicolare a r_1. Si faccia lo stesso per l'altro asse r_2 (detto asse delle ordinate) individuando la perpendicolare a r_2 per P_2. L'intersezione tra le due perpendicolari è proprio il punto P cercato.

In questo modo si è aritmetizzato il piano e molte verifiche possono essere fatte eseguendo calcoli algebrici sulle coordinate dei punti, senza ricorrere a spesso difficili dimostrazioni di geometria sintetica.

L'aver posto delle coordinate nel piano stabilisce anche, infatti, una cor-

rispondenza tra certi sottoinsiemi del piano ed equazioni (o sistemi di equazioni) in due variabili.

In quest'ottica, P appartiene al luogo geometrico $\mathcal{S}$ se sono soddisfatte le oppotune condizioni e ciò accade se e solo se le coordinate di P soddisfano una o più equazioni in due variabili, che verranno dette le equazioni di $\mathcal{S}$.

Prima di fare un esempio richiamiamo la formula della distanza tra due punti nel piano.

Proposizione 9.1.1 *Siano $A = (x_1, y_1)$ e $B = (x_2, y_2)$ due punti del piano ove sia stato fissato un sistema di riferimento cartesiano ortogonale monometrico. Allora $d(A, B)$, la lunghezza del segmento AB, è data dalla espressione:*

$$d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}\,.$$

Questa formula si dimostra applicando il teorema di Pitagora.

Esempio 9.1.2 *Il cerchio γ di centro C e raggio AB (con A, B, C punti del piano) è il luogo dei punti P del piano tali che $d(P, C) = d(A, B)$.*

Quindi verificare che un punto P appartiene a γ consiste nel dimostrare in via sintetica che la lunghezza dal segmento PC è uguale alla lunghezza del segmento AB.

In coordinate, se $C = (x_0, y_0)$ e $d(A, B) = r$, si ha che:

$$\begin{aligned} P = (x, y) \in \gamma \quad &\Leftrightarrow \quad d(P, C) = r \\ &\Leftrightarrow \quad \sqrt{(x - x_0)^2 + (y - y_0)^2} = r \\ &\Leftrightarrow \quad (x - x_0)^2 + (y - y_0)^2 = r^2 \\ &\Leftrightarrow \quad x^2 + y^2 - 2x_0x - 2y_0y + x_0^2 + y_0^2 - r^2 = 0\,. \end{aligned}$$

Questa equazione di secondo grado in due variabili è detta equazione di γ e un punto appartiene a γ se e solo se le sue coordinate verificano tale equazione.

Se $C = (2, 1)$ e $r = 2$, l'equazione di γ diventa:

$$x^2 + y^2 - 4x - 2y + 1 = 0\,.$$

Per verificare se il punto $P = (0, 1)$ appartiene a γ non occorre dimostrare in qualche modo che $d(P, C) = 2$ ma basterà sostituire 0 e 1 rispettivamente alla x e alla y della equazione. Poiché otteniamo $0 + 1 - 0 - 2 + 1 = 0$, possiamo concludere che $P \in \gamma$.

A questo punto lo studio prenderà due diverse direzioni. Da una parte, dati dei luoghi geometrici, alcuni noti fin dall' antichità, vi sarà il problema di scriverne l'equazione.

Dall'altra, problema ben più arduo, data una certa equazione in due variabili, determinare le proprietà del sottoinsieme del piano che tale equazione descrive.

Osservazione 9.1.3 *È facile inoltre verificare che per descrivere l'intersezione tra due luoghi bisogna risolvere il sistema tra le equazioni dei due luoghi, in quanto un punto P appartiene a tale intersezione se e solo se P appartiene ad entrambi i luoghi e ciò accade se e solo se le coordinate di P verificano entrambe le equazioni ossia se sono soluzioni del sistema delle due equazioni.*

Concludiamo questo paragrafo, richiamando un'altra formula di geometria analitica elementare. (Non è che c'entri molto in questo contesto, ma non sapevamo dove inserirla).

Proposizione 9.1.4 *Dati $A = (x_1, y_1)$ e $B = (x_2, y_2)$, punti del piano, il punto medio M del segmento AB ha coordinate:*

$$M = \left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2} \right).$$

9.2 Rette nel piano

Supporremo d'ora in avanti, omettendo a volte di ricordarlo, che nel piano sia stato fissato un sistema di riferimento cartesiano ortogonale monometrico.

Prendiamo in considerazione la più semplice equazione in due variabili:

$$ax + by + c = 0 \tag{9.1}$$

con $a, b, c \in \mathbb{R}$ e $(a, b) \neq (0, 0)$.

Proposizione 9.2.1 *Il luogo dei punti del piano le cui coordinate verificano una equazione di primo grado del tipo (9.1) è una retta.*

Viceversa, data una retta r nel piano, esiste sempre una equazione del tipo (9.1) che è verificata dalle coordinate di tutti e soli i punti di r.

Omettiamo per il momento la dimostrazione-verifica di questo enunciato riservandoci di farla per una proposizione analoga riguardante i piani nello spazio. Quella dimostrazione, ripetuta pedissequamente, mutatis mutandis, si adatta perfettamente a questa proposizione.

Consideriamo l'equazione $ax + by + c = 0$. Essa descrive una retta r. Se $b = 0$, necessariamente $a \neq 0$ ed essa diviene $x = -c/a$. Questa rappresenta l'insieme di tutti i punti aventi la stessa proiezione sull'asse delle ascisse data dal punto P di coordinate $(-c/a, 0)$ e perciò rappresenta la retta verticale, cioè parallela all'asse delle ordinate, passante per tale P.

Se $b \neq 0$ allora possiamo scrivere:

$$y = mx + n\,, \text{ dove } m = -a/b \text{ e } n = -c/b\,;$$

m viene detto *coefficiente angolare* della retta r e n *intercetta* di r.

Osservazione 9.2.2 *Notiamo che la corrispondenza tra luoghi del piano ed equazioni non è biunivoca. Infatti se moltiplichiamo una equazione per un numero reale non nullo, la nuova equazione è verificata dagli stessi valori e pertanto descrive lo stesso luogo.*

n è l'ordinata del punto di intersezione tra la retta r e l'asse delle ordinate.

Si può facilmente verificare che $m = \tan\theta$ dove θ è l'angolo formato tra la retta data r e il semiasse positivo delle ascisse.

Osservazione 9.2.3 *Se r è parallela a r', gli angoli θ e θ, formati rispettivamente da r e r' con il semiasse positivo delle ascisse, essendo angoli corrispondenti di due rette parallele tagliate da una trasversale, sono uguali.*

Possiamo allora dire:

Proposizione 9.2.4 *Se r è la retta di equazione $y = mx + n$ e r' è la retta di equazione $y = m'x + n'$, allora r e r' sono parallele se e solo se $m = m'$.*

Dim. Dimostriamo questo fatto utilizzando non il procedimento geometrico suggerito nell'Osservazione 9.2.3 ma la teoria dei sistemi lineari.

Due rette nel piano sono parallele se e solo se la loro intersezione è vuota oppurre le due rette sono coincidenti. Per l'Osservazione 9.1.3, le coordinate degli eventuali punti appartenenti a $r \cap r'$ sono tutte e sole le soluzioni del sistema:

$$\begin{cases} mx - y = -n\,, \\ m'x - y = -n'\,. \end{cases}$$

La matrice dei coefficienti di questo sistema è:

$$A = \begin{pmatrix} m & -1 \\ m' & -1 \end{pmatrix} .$$

Se $\det A \neq 0$, per il teorema di Cramer, il sistema avrebbe una unica soluzione e quindi le due rette sarebbero incidenti.

Se $\det A = 0$, invece, il sistema non avrebbe soluzioni (se $\det \begin{pmatrix} -1 & n \\ -1 & n' \end{pmatrix} \neq 0$) oppure ne avrebbe infinite (se $\det \begin{pmatrix} -1 & n \\ -1 & n' \end{pmatrix} = 0$). In entrambi i casi r e r' sarebbero parallele: distinte (se $n \neq n'$) oppure coincidenti (se $n = n'$).

Quindi r e r' sono parallele se e solo se $\det A = 0$ e quindi se e solo se $m = m'$. □

Proposizione 9.2.5 *Siano date due rette r e r', rispettivamente di equazioni: $y = mx + n$ e $y = m'x + n'$. Se $m = \tan\theta$ e $m' = \tan\theta'$, allora l'angolo formato da r e r' sarà:*

$$\omega = \theta - \theta' \text{ con } \omega = \arctan \frac{m - m'}{1 + mm'} .$$

Dim. Basta ricordare che $\tan\omega = \tan(\theta - \theta') = \frac{\tan\theta - \tan\theta'}{1 + \tan\theta\tan\theta'}$. □

In particolare, r sarà perpendicolare a r' se si annulla il denominatore della espressione precedente. Abbiamo dunque:

Proposizione 9.2.6 *Siano date due rette r e r' rispettivamente di equazioni $y = mx + n$ e $y = m'x + n'$. Allora r e r' sono perpendicolari se e solo se:*

$$mm' = -1 .$$

Vogliamo ora elencare alcuni casi, ricorrenti nelle applicazioni, in cui si debba ricavare l'equazione di una retta che soddisfi le proprietà richieste dal problema.

9.2.1 Retta per due punti

Siano $A = (x_1, y_1)$ e $B = (x_2, y_2)$ due punti distinti del piano. La retta di equazione:

$$(y_2 - y_1)(x - x_1) = (x_2 - x_1)(y - y_1)$$

è una retta la cui equazione è soddisfatta sia dalle coordinate di A che da quelle di B e quindi è proprio la retta che passa per A e B.

9.2.2 Retta per un punto parallela ad una retta

Sia $A = (x_0, y_0)$ e sia r la retta di equazione $ax + by + c = 0$. La retta $a(x - x_0) + b(y - y_0) + c = 0$ passa per A ed ha lo stesso parametro angolare di r.

9.2.3 Retta per un punto perpendicolare ad una retta

Sia $A = (x_0, y_0)$ e sia r la retta di equazione $ax + by + c = 0$. La retta $b(x - x_0) - a(y - y_0) + c = 0$ passa per A ed ha come parametro angolare b/a, mentre r aveva $m = -a/b$, quindi le due rette sono perpendicolari.

9.2.4 Distanza punto-retta

Sia $A = (x_0, y_0)$ e sia r la retta di equazione $ax + by + c = 0$. Sia H il punto di intersezione tra r e la retta perpendicolare a r passante per A. Allora la distanza $d(A, r)$ tra A e r è la lunghezza del segmento AH. Analiticamente, eseguendo i calcoli, si ha:

$$d(A, r) = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}\,.$$

9.2.5 Bisettrici di due rette

Siano r e r' rette di equazione rispettivamente $ax+by+c = 0$ e $a'x+b'y+c' = 0$. Le bisettrici di r e r' sono il luogo dei punti del piano $P = (x, y)$ tali che $d(P, r) = d(P, r')$. Dovrà allora essere:

$$d(P, r) = \frac{|ax + by + c|}{\sqrt{a^2 + b^2}} = \frac{|a'x + b'y + c'|}{\sqrt{a'^2 + b'^2}} = d(P, r')\,.$$

Questa uguaglianza tra valori assoluti (poiché $|\lambda| = |\mu|$ se e solo se $\lambda = \mu$ oppure $\lambda = -\mu$), diventa:

$$\frac{|ax + by + c|}{\sqrt{a^2 + b^2}} = \frac{|a'x + b'y + c'|}{\sqrt{a'^2 + b'^2}}$$

oppure

$$\frac{|ax + by + c|}{\sqrt{a^2 + b^2}} = -\frac{|a'x + b'y + c'|}{\sqrt{a'^2 + b'^2}}\,.$$

Facciamo un esempio. Sia r la retta di equazione $x+y+1=0$ e r' la retta di equazione $2x-1=0$. Allora le due bisettrici saranno la retta b_1 di equazione:

$$\frac{x+y+1}{\sqrt{2}} = \frac{2x-1}{2} \quad \text{ossia} \quad (2-\sqrt{2})x - \sqrt{2}y - (\sqrt{2}+1) = 0$$

e la retta b_2 di equazione:

$$\frac{x+y+1}{\sqrt{2}} = -\frac{2x-1}{2} \quad \text{ossia} \quad (2+\sqrt{2})x + \sqrt{2}y + (\sqrt{2}-1) = 0\,.$$

9.3 Coordinate cartesiane nello spazio

È possibile fissare nello spazio tre rette perpendicolari tra loro a due a due e incidenti nello stesso punto O. Preso tale punto come origine e prese delle coordinate cartesiane su ciascuna delle tre rette, procedendo come si è fatto all'inizio di §9.1, risulta assegnata ad ogni punto dello spazio una terna di numeri reali, che saranno detti le coordinate di P; questa corrispondenza tra lo spazio ordinario e $\mathbb{R}^3$ è così ottenuta: si proietti ortogonalmente P su ciascuno dei tre assi r_1, r_2, r_3 e sia P_i la proiezione di P su r_i; a ciascun P_i è associato un numero reale x_i; la terna di numeri reali (x_1, x_2, x_3) sono le coordinate di P nel sistema di riferimento cartesiano ortogonale monometrico fissato.

Che la corrispondenza sia biunivoca lo si dimostra analogamente a come fatto per le coordinate cartesiane del piano.

Assegnamo ai tre assi cartesiani i nomi di asse delle ascisse (o asse delle x), asse delle ordinate (o asse delle y) e asse delle altezze (o asse delle z).

Tutte le considerazioni fatte per le coordinate nel piano restano ancora valide, in particolare si ha la possibilità di risolvere con calcoli analitici problemi di natura geometrica.

Proposizione 9.3.1 *Siano A e B due punti dello spazio di coordinate rispettivamente (x_1, y_1, z_1) e (x_2, y_2, z_3). Se $d(A, B)$ è la distanza tra A e B, si ha:*

$$d(A,B) = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2}$$

Proposizione 9.3.2 *Siano A e B due punti dello spazio di coordinate rispettivamente (x_1, y_1, z_1) e (x_2, y_2, z_2). Se M è il punto medio del segmento AB, M ha le seguenti coordinate:*

$$M = \left(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2}, \frac{z_1+z_2}{2}\right)\,.$$

9.4 Vettori geometrici nello spazio

L'introduzione di coordinate cartesiane nel piano o nello spazio permette di effettuare una utilissima ulteriore identificazione.

Ogni punto dello spazio (rispettivamente del piano) può essere pensato come l'estremo libero di un vettore geometrico applicato nell'origine O del sistema di riferimento (vedi Esempio 5.1.3). Alla terna di numeri reali (a, b, c) possiamo quindi far corrispondere, oltre che il punto dello spazio $P = (a, b, c)$, anche il vettore geometrico $\overrightarrow{OP}$.

Siano $A_1 = (1, 0, 0), A_2 = (0, 1, 0), A_3 = (0, 0, 1)$ e poniamo:

$$\mathbf{i} = \overrightarrow{OA}_1\,, \quad \mathbf{j} = \overrightarrow{OA}_2\,, \quad \mathbf{k} = \overrightarrow{OA}_3\,.$$

I vettori $\mathbf{i}, \mathbf{j}, \mathbf{k}$, costituiscono una base per lo spazio vettoriale V_O^3 dei vettori applicati nel punto O e, se pensati come vettori geometrici liberi, cioè come rappresentanti della loro classe di segmenti orientati equivalenti, sono anche una base dello spazio dei vettori geometrici liberi dello spazio.

La corrispondenza indicata è un isomorfismo tra gli spazi vettoriali V_O^3 e $\mathbb{R}^3$. Lasciamo per esercizio la verifica di queste affermazioni.

Questo isomorfismo ci permette di effettuare le seguenti osservazioni che sarà molto utile tenere presenti.

Osservazione 9.4.1 *1) Sia dato $v \in V_O^3, v \neq 0$, sia $\mathbb{R}v$ il sottospazio di V_O^3 costituito dai multipli di v, $\mathbb{R}v = \{w \in V_O^3 \,|\, w = \lambda v, \lambda \in \mathbb{R}\}$.*

Se v ha coordinate (a, b, c), cioè $v = \overrightarrow{OA}$, con $A = (a, b, c)$, allora v corrisponde alla retta passante per l'origine e per il punto A.
2) Se $v = \overrightarrow{OA}$ ha coordinate (a, b, c) e se $|v|$ indica la lunghezza del segmento OA, allora $|v| = \sqrt{a^+ b^2 + c^2}$. Infatti questa è la distanza di $A = (a, b, c)$ dall'origine O.
3) Consideriamo la seguente funzione

$$<, >: V_O^3 \times V_O^3 \to \mathbb{R}\,, \quad < v, w >= |v||w| \cos\theta\,,$$

dove θ è l'angolo formato dai segmenti $OA = v$ e $OB = w$.

Si può dimostrare che $<, >$ è un prodotto scalare definito positivo (utilizzando il teorema di Carnot sui triangoli e le formule di prostaferesi). Inoltre la base $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ è una base ortonormale per $<, >$. Allora si può provare che, se (a, b, c) e (a', b', c') sono le coordinate dei vettori v e w, rispettivamente, si

ha che:

$$< v, w >= (a, b, c) \cdot \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = aa' + bb' + cc' ,$$

cioè eseguire il prodotto scalare $< v, w >$ dà lo stesso risultato che fare il prodotto scalare canonico di $\mathbb{R}^3$ con le coordinate di v e w. Per dimostrare ciò è sufficiente verificarlo per i vettori della base $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$.
4) Due vettori v, w sono perpendicolari se e solo se $< v, w >= 0$. Infatti, $< v, w >= 0$ se e solo se l'angolo formato dai due vettori ha coseno nullo, cioè è un angolo retto.
5) Dati due vettori v, w, il numero $\frac{<v,w>}{|v|}$ indica la lunghezza della proiezione di w sulla retta $\mathbb{R}v$, lunghezza presa con segno positivo se la proiezione di w appartiene alla stessa semiretta che contiene l'estremo libero di v, mentre è presa con segno negativo se la proiezione di w appartiene alla semiretta opposta.

Definizione 9.4.2 *Consideriamo l'operazione $\Lambda : V_O^3 \times V_O^3 \to V_O^3$ così definita. Siano u e v due vettori dello spazio con coordinate rispettivamente (x_1, y_1, z_1) e (x_2, y_2, z_2). Allora $w = u \wedge w$ ha coordinate (x_3, y_3, z_3) con:*

$$x_3 = \det \begin{pmatrix} y_1 & y_2 \\ z_1 & z_2 \end{pmatrix} , \quad y_3 = -\det \begin{pmatrix} x_1 & x_2 \\ z_1 & z_2 \end{pmatrix} , \quad z_3 = det \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} .$$

Il vettore $w = u \wedge v$ viene detto il prodotto vettore *di u e v.*

Si dimostra che la definizione è ben posta, non dipende cioè dalla scelta delle coordinate.

Proposizione 9.4.3 *1) L'operazione Λ è bilineare.*
2) Λ è anticommutativa, cioè $u \wedge v = -v \wedge u$.
3) $u \wedge v = 0$ se e solo se u e v sono linearmente dipendenti.
4) $< u \wedge v, w >= \det \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix}$, dove (z_1, z_2, z_3) sono le coordinate di w.
5) $u \wedge v$ è perpendicolare a entrambi i vettori u e v.

La dimostrazione di questi fatti è immediata se si osserva che, per $j = 1, 2, 3$, la j-esima coordinata di $u \wedge v$ è il determinante del minore ottenuta cancellando dalla matrice $\begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \\ z_1 & z_2 \end{pmatrix}$ la j-esima riga, moltiplicato per $(-1)^{j+1}$.

Il prodotto vettore si utilizza spesso per determinare rapidamente un vettore perpendicolare a due vettori dati ed inoltre è di notevole importanza nella teoria delle grandezze fisiche vettoriali.

9.5 Piani nello spazio

Se v è un vettore nello spazio, non nullo, come abbiamo osservato nell'Osservazione 9.4.1, il numero $\frac{<v,w>}{|v|}$ indica la lunghezza, con segno, della proiezione di w sulla retta v.

Allora è chiaro che, dati due vettori u e w, si ha

$$\frac{< v, w >}{|v|} = \frac{< v, u >}{|v|}$$

se e solo se la proiezione degli estremi liberi di u e w sulla retta v coincidono.

Sia Q un punto appartenente alla retta $\mathbb{R}v$. Dalla geometria elementare sappiamo che:

$$\pi = \{P \,|\, Q \text{ è la proiezione ortogonale di } P \text{ su } v\}$$

è un piano, il piano passante per Q e perpendicolare e $\mathbb{R}v$. Allora:

$$P \in \pi \Leftrightarrow \frac{< v, \overrightarrow{OP} >}{|v|} = \frac{< v, \overrightarrow{OQ} >}{|v|} \Leftrightarrow < v, \overrightarrow{OP} >=< v, \overrightarrow{OQ} > \ .$$

Supponiamo che: $v = (a, b, c)$, $\overrightarrow{OP} = (x, y, z)$ e $< v, \overrightarrow{OQ} >= d$.

(Utilizziamo liberamente le coordinate per i vettori geometrici oltre che per i punti dello spazio. Un minimo di attenzione permette, facendo di volta in volta mente locale, di evitare le confusioni.)

Poiché il prodotto scalare $<, >$ coincide, quando si opera in coordinate, col prodotto scalare canonico di $\mathbb{R}^3$, si ha che:

$$P \in \pi \Leftrightarrow ax + by + cz = d\,.$$

Abbiamo così provato il seguente enunciato:

Proposizione 9.5.1 *Se* $(a, b, c) \neq (0, 0, 0)$*, l'insieme dei punti* P *dello spazio le cui coordinate verificano una equazione di primo grado del tipo* $ax + by + cz = d$*, è un piano.*

Viceversa, se π è un piano, sia $v \neq 0$ un vettore tale che la retta $\mathbb{R}v$ sia la retta per l'origine perpendicolare a π.

Sia Q il punto di intersezione tra v e π. Allora: $P \in \pi \Leftrightarrow < v, \overrightarrow{OP} > = < v, \overrightarrow{OQ} >$. Quindi esiste $(a, b, c) \neq (0, 0, 0)$, le coordinate di v, e $d = < v, \overrightarrow{OQ} >$ tali che, se $P = (x, y, z)$,

$$P \in \pi \Leftrightarrow ax + by + cz = d\,.$$

Quindi ogni piano ammette una equazione del tipo $ax + by + cz = d$.

Tale corrispondenza tra piani ed equazioni non è biunivoca ma, come già notato nell'Osservazione 9.2.2, l'equazione di un piano è definita a meno di una costante moltiplicativa non nulla.

Osservazione 9.5.2 *Un piano, di equazione* $ax + by + cz = d$*, passa per l'origine, se e solo se* $d = 0$*.*

Definizione 9.5.3 *I numeri della terna* (a, b, c) *vengono detti* parametri direttori *del piano* π *di equazione* $ax+by+cz = d$*. Essi individuano un vettore perpendicolare al piano* π*.*

Come fatto per le rette nel piano, vogliamo elencare anche per i piani nello spazio alcuni casi frequenti in cui si debba ricavare l'equazione di un piano che soddisfi le proprietà assegnate.

Proposizione 9.5.4 (Piano per tre punti assegnati) *Dati nello spazio tre punti non allineati,* $A = (x_1, y_1, z_1)$*,* $B = (x_2, y_2, z_2)$ *e* $C = (x_3, y_3, z_3)$*, la seguente equazione (espressa in effetti in modo insolito):*

$$\det \begin{pmatrix} x - x_1 & x_2 - x_1 & x_3 - x_1 \\ y - y_1 & y_2 - y_1 & y_3 - y_1 \\ z - z_1 & z_2 - z_1 & z_3 - z_1 \end{pmatrix} = 0\,, \tag{9.2}$$

rappresenta il piano passanti per i tre punti dati.

Dim. Sviluppiamo il calcolo del determinante in (9.2):

$$\begin{aligned} (x-x_1)[(y_2-y_1)(z_3-z_1)-(y_3-y_1)(z_2-z_1)] &\ - \\ -(y-y_1)[(x_2-x_1)(z_3-z_1)-(x_3-x_1)(z_2-z_1)] &\ + \\ +(z-z_1)[(x_2-x_1)(y_3-y_1)-(x_3-x_1)(y_2-y_1)] &= 0\,. \end{aligned}$$

Svolgendo ulteriormente i calcoli si ottiene una equazione di primo grado $ax+by+cz=d$ nelle tre variabili x, y, z. Inoltre, sostituendo le coordinate dei punti A, B, C al posto di x, y, z, si trova che:

- sostituendo le coordinate del punto A, la prima colonna della matrice in (9.2) diventa una colonna di zeri, quindi il determinente è nullo e A appartiene al piano individuato;
- sostituendo le coordinate del punto B, la prima e la seconda colonna della matrice in (9.2) diventano uguali e quindi il determinente è nullo e B appartiene al piano individuato;
- stesso discorso per le coordinate di C, le quali rendono uguali la prima e la terza colonna della matrice (9.2).

□

Se $a=b=c=0$, nel caso precedente ciò vuol dire che i tre punti erano allineati e dunque vi sono infiniti piani che li contengono tutti.

Siano $a_1x+b_1y+c_1z=d_1$ e $a_2x+b_2y+c_2z=d_2$ due piani, rispettivamente π_1 e π_2, nello spazio. Vogliamo studiare le condizioni che debbono essere verificate affinché essi siano paralleli. Per far ciò determiniamo la loro intersezione, il che equivale a risolvere il seguente sistema lineare:

$$\begin{cases} a_1x+b_1y+c_1z=d_1\,, \\ a_2x+b_2y+c_2z=d_2\,. \end{cases}$$

Abbiamo che

$$A=\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}$$

è la matrice dei coefficienti del sistema, mentre

$$A'=\begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \end{pmatrix}$$

è la matrice completa.

Se car$A = 2$, per il teorema di Rouché-Capelli, il sistema ammette soluzioni e pertanto i due piani non sono parelleli.

Se invece car$A = 1$ e car$A' = 2$, il sistema non ammette soluzioni e pertanto i due piani sono paralleli e distinti.

Se infine car$A =$ car$A' = 1$, allora vuol dire che le due righe di A' sono proporzionali e quindi l'equazione di π_1 è ottenuta moltiplicando per una costante l'equazione di π_2 e perciò, come luoghi, i due piani coincidono.

Abbiamo quindi provato che:

Proposizione 9.5.5 *Due piani sono paralleli se e solo se hanno i parametri direttori proporzionali.*

Questo risultato ci permette di risolvere facilmente problemi del tipo seguente:

Problema 9.5.6 *Determinare il piano π passante per $A = (x_0, y_0, z_0)$ e parallelo al piano τ di equazione data $ax + by + cz = d$.*

Affinché π sia parallelo a τ i parametri direttori di π devono essere proporzionali a quelli di τ; ad esempio, per comodità possiamo sceglierli uguali. Allora il piano π di equazione:

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

ha i requisiti richiesti: ha gli stessi parametri direttori di τ e quindi è parallelo a τ; le coordinate di A verificano l'equazione di π e perciò $A \in \pi$.

Definizione 9.5.7 *L'angolo fra due piani è l'angolo formato dai vettori geometrici aventi come coordinate i parametri direttori dei due piani.*

Quindi se π_1 ha equazione $a_1x + b_1y + c_1z = d_1$ e π_2 ha equazione $a_2x + b_2y + c_2z = d_2$, siano v_1 e v_2 i vettori aventi come coordinate rispettivamente (a_1, b_1, c_1) e (a_2, b_2, c_2).

L'angolo θ formato da π_1 e π_2 è tale che:

$$\cos\theta = \frac{< v_1, v_2 >}{|v_1|\,|v_2|} = \frac{a_1a_2 + b_1b_2 + c_1c_2}{\sqrt{a_1^2 + b_1^2 + c_1^2}\,\sqrt{a_2^2 + b_2^2 + c_2^2}}\,.$$